Structure and Dynamics of High Mobility Group Box Protein 1 (HMGB1) Complexed With Inhibitors of Its Proinflammatory Activity: A Combined Nuclear Magnetic Resonance and Computational Chemistry Approach

Thesis

UNRESTRICTED

Luca Mollica

Structure and Dynamics of High Mobility Group Box Protein 1 (HMGB1) complexed with Inhibitors of its Proinflammatory Activity: A Combined Nuclear Magnetic Resonance and Computational Chemistry Approach

Degree of Doctor of Philosophy

in

Cellular and Molecular Biology

The Open University

September 2008

Sponsoring Establishment:

DIBIT, University Vita-Salute San Raffaele

Milan, Italy

Submission date: 20 Aug. 2008
Date of award: 20 Nov. 2008

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Table of Contents

Abstract

High-mobility group box 1 protein (HMGB1) is a nuclear component, but extracellularly it serves as a signaling molecule involved in acute and chronic inflammation, for example in sepsis and arthritis. The identification of HMGB1 inhibitors is therefore of significant experimental and clinical interest. We show that glycyrrhizin, a natural anti-inflammatory and antiviral triterpene in clinical use, inhibits HMGB1 chemoattractant and mitogenic activities, and has a weak inhibitory effect on its intranuclear DNA binding function. NMR and fluorescence studies indicate that glycyrrhizin (and its analogue carbenoxolone) binds directly to HMGB1 (K_d ~ 150 μM), interacting with two shallow concave surfaces formed by the two arms of both HMG boxes, altering the internal mobility of proteins (on the millisecond timescale for Box A, on the nanosecond timescale for Box B) without significantly altering their overall diffusive properties. Due to the reduced number of experimentally determined intermolecular contacts, by means of two different computational approaches (QRFF for the Box A - carbenoxolone complex, HADDOCK data driven docking for the glycyrrhizin containing complexes) it was possible to obtain an atomic level model of the protein-ligand interactions.

The presented results explain in part the anti-inflammatory properties of glycyrrhizin, and might direct the design of new derivatives with improved HMGB1-binding properties.

List of Abbreviations, Acronyms and Definitions

CBNX	Carbenoxolone
CNS	Central nervous system
COPD	Chronic Obstructive Pulmonary Disease
COX	Cyclooxygenase
CPMD	Carr-Parrinello Molecular Dynamics
δ	Chemical shift
$D_{x,y,z}$	Diffusion tensor axis
DFT	Density Functional Theory
FF	Force Field
FL	Full Length
fMLP	formyl Met-Leu-Pro
FRAP	Fluorescence Recovery After Photobleaching
GC	Glucocorticoids
GL	Glycyrrhizin
GR	Glucocorticoid Receptors
HAC	Histone Acetylase
HDAC	Histone Deacetylase
hetNOE	Heteronuclear Overhauser Effect
HMBC	Heteronuclear Multiple Bond Quantum Coherence
HMGB1	High Mobility Group B1
HSQC	Heteronuclear Single Quantum Coherence
IKK	Ik Beta Kinase
IL	Interleukin
J	Spin-spin Coupling constant
LOX	Lipooxigenase
MAPK	Mitogen Activated Protein Kinase
MD	Molecular Dynamics

MM	Molecular Mechanics
MMP	Matrix Metalloprotease
NfkB	Nuclear Factor kB
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Enhancement
NOESY	Nuclear Overhauser Enhancement Spectroscopy
NSAID	Non Steroidal Antiinflammatory Drug
PLA	Phospholipase
PME	Particle Mesh Ewald
POPC	Phosphatidylcholine
POPE	Phosphatidylethanolamine
RMSD	Root Mean Square Displacement
ROESY	Rotating frame Spectroscopy
TNF	Tumor Necrosis Factor
TOCSY	Total Correlation Spectroscopy
RDC	Residual Dipolar Couplings
SAID	Steroidal Antiinflammatory Drug
SMC	Smooth Muscle Cells
STD	Saturation Transfer Difference
τ_c	Rotational correlation time
T_1	Longitudinal relaxation time
T_2	Transversal relaxation time
TLR	Toll-Like Receptor

1. Introduction

1.1 Safer drugs, saving money: new routes to drug discovery.

Drug discovery is a process that covers different aspects of biomedical sciences, from basic knowledge to economical and marketing management, and that has a central role in the economic systems of all the richest countries worldwide: it is very long in terms of years of investment of human and economical resources as well. It has been estimated that the complete route that goes from a simple molecule to the pharmaceutical market covers approximately 15 years including *in silico*, *in vitro* and *in vivo* experimental procedures (http://www.linnaeus.uu.se/online/pharm/lakemedelutveckling.html). Not all drugs have a successful life, that means that not all the molecules investigated by research centres and/or companies come to marketplace: during 1980 to 1992, the numbers of applications to investigate potential new drugs and numbers of new chemical entities approved each year by the FDA has remained flat at approximately 90 and 20, respectively (Kaitin et al., 1993). Economically speaking, current estimates for successful launch of a single new medication are in between US $300 million and US $800 million, nearly an order of magnitude higher than that for the NIH Molecular Libraries initiative for publicly available libraries (Di Masi et al., 1992; Rawlins, 2004).

In this century infant mortality and big infectious killers have been progressively replaced by cardiovascular diseases and cancer as the major causes of death. Decade upon decade, life expectancy has increased and, with it, the health care requirements of the elderly. A victim of the success of health care most broadly defined, the target for new pharmacological intervention has moved and is getting progressively more difficult to hit. In an analysis of 55 new chemical entities approved in the US in 1987 to 1989, 9 (16%) were classified by the FDA as important therapeutic gain, 15 (27%) as modest gain and 29 (53%) as little or no gain (Kaitin et al., 1991). A repeat analysis covering 1990 to 1992 found similar proportions (Kaitin et al., 1994). Therefore, in the last ten years in particular many drugs have been created as "me-too" drugs (Mason et al., 1998), that are compounds that are very similar to already patented molecules but that

have been slightly modified in order to avoid, *e.g.*, side effects and because they represent safe markets. If this strategy avoided the investment of the great capitals that characterized the "Big Pharma" business between 70s and 80s, on the other side revealed a crisis within this world in terms of lack of novel scientific relevant discoveries in this field. Many problems must be faced during the drug discovery process, ranging from quantum chemistry studies of the molecules under investigation to the marketing strategies: the most relevant features of this process from a chemical and biological point of view are the the ability of being delivered in the organism (usually faced during the clinical phase I) and the evaluation of side effects. This last point should usually be covered during the phase III in principle (Di Masi et al., 1992), but they emerge during the market life of the drug and "me-too" drugs are designed in order to minimize the side effects of the original drug.

A way to avoid the great expense by pharmaceutical companies and to avoid any kind of problems with patenting is the usage of old drugs, originally designed with some purposes, for diseases that had not been considered in the past (Chong et al. 2007). The advantage of such an approach is both economic and pharmaceutical: the clinical trials can be largely or totally avoided, the side effects for old drugs are very well known and in many cases patents are no longer valid. In the last decade many drugs have been reused in the treatment of pathologies different than the ones for which they have originally been thought: for example, Nifepidine, originally designed to be a blood pressure control drug, has been successfully used for the treatment of hemochromatosis (Ludwiczek et al. 2007), or Astemizole was transferred from allergic rhinitis to malaria cure (Chong 2006). In particular, neurology and neurosciences are among the branches that gained major benefit from this strategy: Eritropoietin (EPO), a well known anemia drug (and sport dopant), has demonstrated its efficiency in protecting nerves from oxygen deficiency in brain ischemia (Lykissas et al., 2008). Also, multiple sclerosis and Huntington disease in the last year have obtained large benefits from the rediscovery of two drugs like, respectively, Atorvastatin (Li et al. 2007) and Clioquinol (Gouras et al., 2001).

A family of drugs that are suitable for this rediscovery are anti-inflammatory

drugs (Rhen et al., 2005; Rosen et al. 2005; Maroon et al., 2006): nowadays both steroidal and nonsteroidal drugs side effects are considered too risky for human diseases, hence some efforts have to be considered in order to create novel kinds of anti-inflammatory drugs. Non steroidal anti-inflammatory drugs (NSAID) regulate the inhibition of the prostaglandine synthesis, being able to interact with both cyclooxygenase (COX) enzymes at the same time or with COX of type 2 only, whereas steroidal drugs (SAID) are able to interfere with NF-kappaB signalling that after activation is able to make the cells release cytokines and late mediators of inflammation. In both cases, side effects can be very dangerous or unhealthy: many cases of gastrointestinal, renal or vascular failures have been recorded in the recent past in patients treated with NSAIDs, whereas a high dosage of SAIDS induces inhibition of longitudinal growth in children, hypertension and osteoporosis in adults (van der Eerden, 2003; Ullian et al., 1999). For these reasons, the ability of finding a new way of thinking to anti-inflammatory drugs is crucial for avoiding high risks in inflammatory diseases. Unfortunately, the possibility of ameliorating the modes of action of COX-related drugs is very low due to the differential expression of COX-1 and COX-2 and to their nature (see further): in this sense in the last years the majority of efforts have been focused on alternative drugs that could interfere with the NF-kappaB pathway or in the discovery of new inflammation pathways.

1.2 Steroidal and non steroidal anti-inflammatory drugs

The two most well estabilished ways of signalling inflammation are the glucocorticoid receptor (GR) way (Rhen et al, 2005; Maroon et al., 2006) and the ecosanoids production regulated by means of arachidonic acid pathway (Khanapure et al. 2007). The drugs that are involved in the blockage of the first pathway are chemically derivatives of cortisol and hence are called steroidal anti-inflammatory drugs (SAID) and are simply usually indicated as glucocorticoids (GC), whereas for blocking Arachidonic acid cascade drugs that are able to inhibit the activity cyclooxygenase 1 and 2 (COX1 and COX2) are used that are called, due to their chemical nature, non-steroidal anti-inflammatory drugs (NSAID). In the two following subsections, the modes of action

of these two distinct types of drugs are highlighted together with a sketch of the biochemical mechanism in which they are involved.

1.2.1 Glucocorticoids

Glucocorticoids are a class of steroid hormones originally identified to be crucial in controlling the blood glucose content, from which the name derived. Glucocorticoids must be distinguished from mineralocorticoids and sex steroids mainly because of receptors involved in their biochemistry and/or cells type involved. The glucocorticoid receptor (GR) is a member of the steroid-hormone-receptor proteins family (Breuner et al., 2002), it is expressed almost in all type of cells, and its expression is mainly regulated by proinflammatory agents tumor necrosis factor (TNF) alpha and interleukin (IL) 1. Its endogenous ligand is cortisol, which is secreted by adrenal cortex and converted to cortisone by 11β-hydroxysteroid dehydrogenase II (Sahia, 2008) in response to neural, cytokine and endocrine stimuli (see Fig.1.1a). Three different modes of action have been identified for cortisol (Fig.1.2): i. the complex formed by cortisol and GR is recruited by the nucleus for activating gene expression directly interacting with DNA in a sequence specific fashion (GR responsive elements); ii. no direct interaction between DNA and cortisol-receptor occurs, but the interaction is mediated by various nuclear factors (see further) that are able to regulate gene expression; iii. other membrane receptors are activated by the formation of ligand-receptor complex (non genomic pathway). Both in the genomic and in the non-genomic ways phosphorylation is the most relevant posttranslational modification that is performed by Mitogen Activated Protein Kinase (MAPK) mainly in a ligand dependent manner (Ismaili et al., 2004): GR is also frequently ubiquitinated and sumoilated for targeting proteosomal degradation (Holmstroon et al. 2003; Le Drean et al. 2002). The anti-inflammatory activity of cortisol and derivatives is mainly related to the inhibition of the biosynthesis of eicosanoids: this process is controlled by the activation of Annexin I and of MAPK phosphatase 1, and by the downregualation of genes that control the expression level of cyclooxygenase 2. Annexin I is an anti-inflammatory protein that interacts directly with PLA2, the enzyme that converts the membrane phospholipids into arachidonic acid

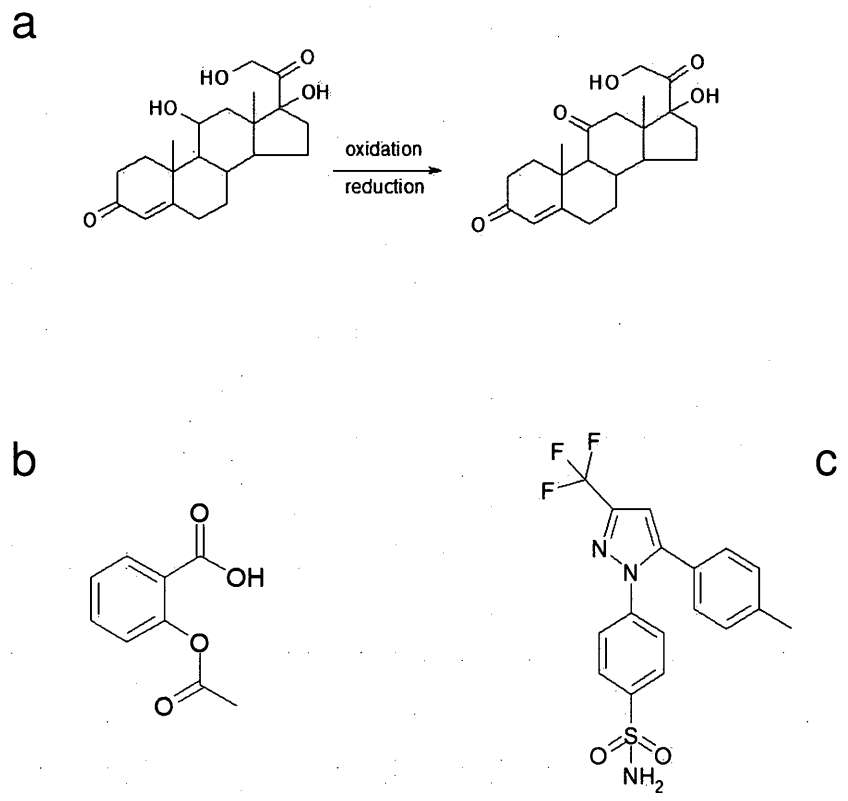


Fig. 1.1. a. Cortisol-cortison interconvertion through an oxidoreduction reaction b. aspirin (acetyl salycilic acid) molecule; c. celecoxib (commercial name: Celebrex) molecule.

(Kim et al., 2001) maintaining the membrane haemostatic balance unaltered; the formation of Annexin I – PLA2 complex is then able to block the release of arachidonic acid. MAPK phosphatase 1 is the protein that controlsth the activation of all the members of MAPK proteins family and of Jun N terminal kinase, hence it's expression/activation is able to inactivate the kinase cascade that is necessary for controlling the enzymatic activity of PLA2 and also to block the phosphorylation of PLA2 itself (Gjion et al.1999). If the above mentioned anti-inflammatory therapy is really efficient and rapid, there are many disadvantages that have to be considered with care (Schacke et al., 2002): hypertension is one of the most evident cortisol and cortisol-derivative side effects, and also wound repair is in some extent inhibited due to the

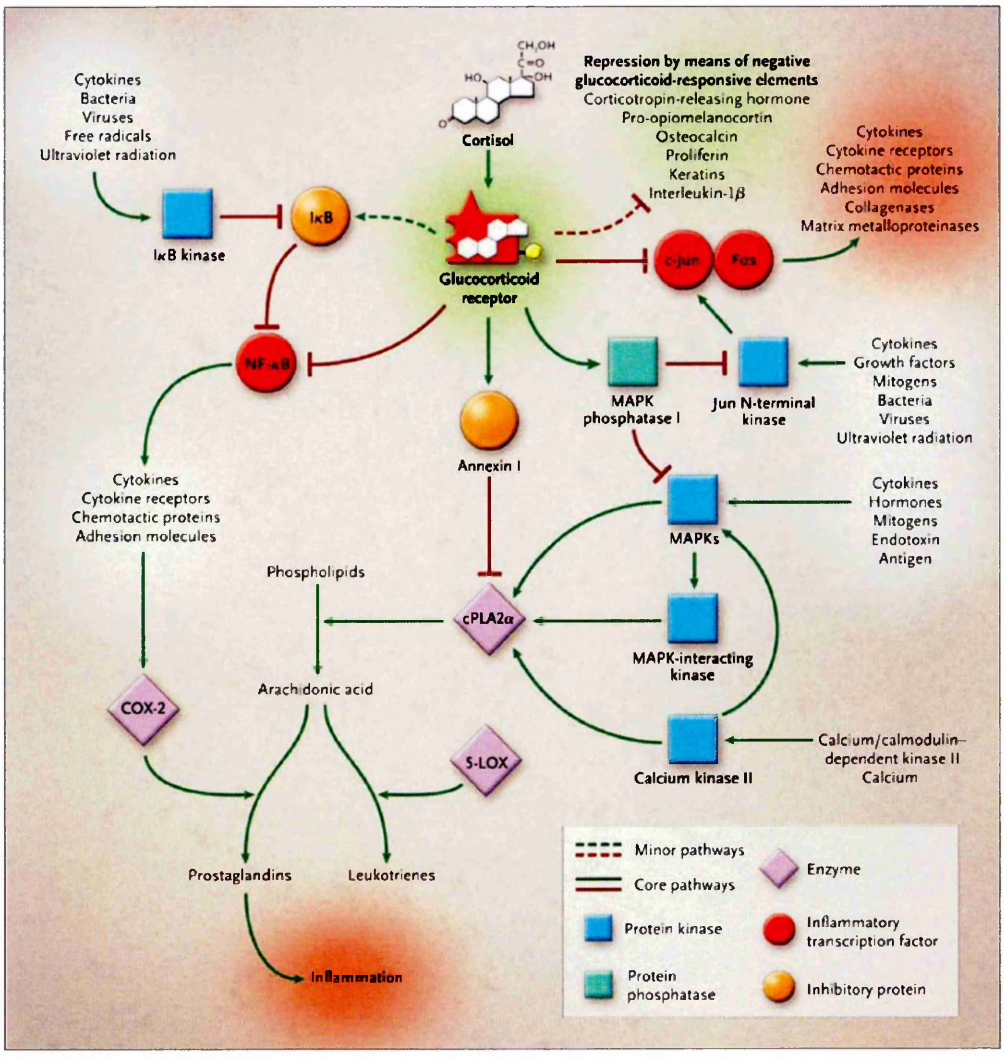


Fig. 1.2 Glucocorticoid pathways during response to inflammation (from Rhen et al, 2005).

downregulation of genes that control the expression level of collagen and matrix metalloproteases (MMPs). Moreover, in children a GC based therapy would have detrimental effect on growth, whereas in elderly people a serious problem facing with osteoporosis has been detected up to date.

1.2.2 Prostaglandins, COX-1, COX-2 and the mode of action of NSAIDs

Traumatic, chemical or autoimmune tissue injury put any type of cells in condition to release short-lived and localized hormones that can induce inflammation, pain, fever once they have been translocated to extracellular matrix and then they are able to interact with respective receptors: these molecules are called prostaglandins (Khanapure et al. 2007). They are the most abundant member of eicosanoids, small molecules all involved in inflammatory processes at different extent, so named because of the chemical origin from the 20 carbon atoms that form the backbone of arachidonic/prostanoic acid (Fig.1.3a): the biological activity of ecoisanoids is various and can be both favourable and unfavourable, ranging from, *e.g.*, checking ordinary functions in the gastrointestinal tract to amplifying pain or restricting vessels diameter during inflammation. Together with prostaglandins, the family of eicosanoids is also composed by thromboxanes , that are exclusively involved in the coagulation cascade, and by leukotrienes, involved in cells stimulation for enhancing the production of interleukins and interferons during inflammatory process and are of central importance in autoimmune diseases (Dharmananda, 2003; Funk, 2001).

The pathway that brings to the synthesis of prostaglandins starts from the cell membrane: after injury Phospholipase 2 (PLA2) recruits phosphatidylethanolamine (POPE) and phosphatidylcoline (POPC) and converts them into Arachidonic acid, which is then converted into eicosainoids by the enzymes lipooxygenase (LOX) and cyclooxygenase (COX): the former produces leukotrienes, the latter thromboxanes and prostaglandins (Fig.1.3b). Between 60s and 70s John Vane and his coworkers were able to find a link between an anti-inflammatory drug of longstanding tradition like aspirin (it was synthesized in 1875, Fig.1.1b), cyclooxygenases and prostaglandins (Vane, 2003): aspirin is able to inhibit the enzymatic activity of COX, hence decreasing the production of proinflammatory eicosanoids and reducing inflammation, pain, fever and blood clotting. Due to the different chemical nature of aspirin (acetylsalicylic acid) with respect to steroidal drugs, aspirin and its derivatives were called non steroidal antiinflammaotry drugs (NSAID): the same name was adopted for drugs that did not display chemically a cortisol-like structure. 20 years later two active forms of

cyclooxygenase were discovered and named COX-1 and COX-2 (Vane et al., 1995).

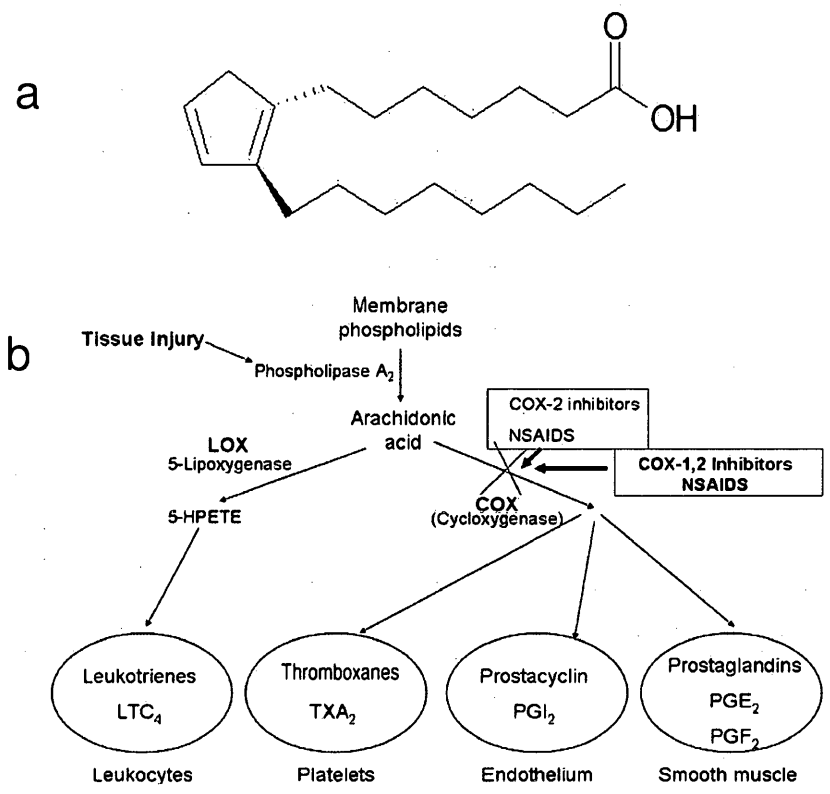


Fig. 1.3 a. Arachidonic acid; b. COX-1 and COX-2 pathways during inflammation (from Maroon et al., 2005).

COX-1 is constitutively expressed almost in all cells and it has a role in balancing homoeostasis and in protecting intestinal mucosa (Seibert et al., 1994), whereas COX-2 is expressed only in cells that are involved in inflammation (Moncado et al., 1976): therefore, COX-2 is the only cyclooxygenase responsible for mediating inflammatory response in cells. The mode of action of aspirin had been already explained by Vain in

70s, but there was no clue for gastrointestinal side effects: with the discovery of COX2 and the comprehension of the cellular and tissue localization of both COXs it was evident that only selectively blocking COX2 was possible to avoid side effects of aspirin, hence a new class of drugs that could selectively inhibit COX2 was the new goal of pharmaceutical industry at the beginning of 90s. This led in 1998 to the approval of celecoxib (Fig.1.1c) by FDA for the treatement of arthritis (Clemett et al., 2000), and within some moths other two drugs of this kind (rofecoxib and valdecoxib) were approved and in some months the three (jergally named "The Big Three") became pharmaceutical market leaders among anti-inflammatory drugs and among drugs in general, leading, *e.g.*, to about 20 million people in US only using selective COX2 inhibitors for medium and long term usage as anti-inflammatory therapy (Maroon et al., 2008). In 2004 the APPROVe study (Hegmann, 2005) lead by the Merck Research Laboratories indicated that after 18 months of usage of rofecoxib (commercially Merck's Vioxx) many patients manifested some kinds of thrombolytic events, even myocardial infarction in the most dramatic cases, and the percentage of risk was even higher in patients with a basal high risk (Caldwell et al., 2006). The reason of this side effect was the breaking of the equilibrium between the unchanged production of thromboxanes by COX-1 in platelets and the inhibition of the production of prostacyclins by COX-2 : the former are thrombogenic, the latter are thrombolitic, hence selective COX-2 inhibitors lead to an increase percentage of risk of thrombolitic events. This, together with other similar studies, withdrew the usage of such drugs and pushed towards the research of novel roads for drug discovery in the field of inflammation and of anti-inflammatory drugs.

1.3 NF-kappaB based inflammation pathways

In the light of the problems presented by both steroidal and nonsteroidal drugs and listed in the previous section, many efforts have been done since 80s in order to discover new pathways that could lead to the design of novel anti-inflammatory drugs. The most promising one was the NF-kappaB pathway, that pushed towards a nuclear view of inflammation with respect to antiCOX based herapies: NF-kappaB has been extensively

studied since mid 80s (Gilmore et al., 1986) and more strongly related to inflammation and oxidative stress at the beginning of 90s (Bonizzi et al., 2004), and this field of research revealed suddenly its importance in oncology due to the strong relationship between certain types of cancer and inflammation with NF-kappaB acting as a bridge between two apparently distinct pathologies.

In practice no efforts have been done in order to design new drugs that could interfere directly with the NF-kappaB pathway because of the very poor way of controlling the fate of such a drug in expressing or silencing unwanted genes in the complex regulatory network. Nevertheless, since 90s many efforts have been done to clarify the effect of natural drugs or existing synthetic drugs originally designed for other purposes on the NF-kappaB pathway in order to understand the possibility of interfering with the inflammation process. The advantage of using natural compounds that are present in common foods is to have diet integrators that could be helpful in some specific situation (*i.e.*, athletes suffering inflammation, high sensibility to drugs) (Maroon et al., 2006) ; the rediscovery of old drugs is extremely advantageous from both an economical point of view and an experimental point of view due to the possibility of avoid clinical phases for monitoring side effects (drugs of long-standing tradition have been already tested and monitored).

In the following subsections NF-kappaB will be introduced in more details, giving some basics of its interactions and its networking among different genes; afterwards, both some natural drugs (curcumin, catechins and resveratrol) and the modes of action of old "recycled" commercial drugs (thalidomide) within the NF-kappaB pathway will be shown.

1.3.1 The NF-kappaB pathway, a nuclear weapon towards inflammation

A search of publications indexed for Medline and concerning NF-kappaB reveals that there are approximately 23000 papers published and almost 10 % of them are reviews about it, indicating that maybe NF-kappaB is the most studied proteins worldwide. Its popularity derives from the discoveries in the first half of 80s of three proteins that displayed a great variability in subcellular localization and in shuttling between the

cytoplasm and the nucleus (Gilmore, 1990): these three proteins had been named NF-kappaB, v-Rel and Dorsal and despite the same kind of subcellular localization they were involved respectively in immunity, oncogenesis and development (Karin, 2006) . In the following years it became clear that apart from the involvement in different tissues and systems the three proteins belonged to the same large family of proteins that are nowadays almost always referred as NF-kappaB, maybe because it is the protein to which more effort has been dedicated in these years and because the immunological studies were the ones that provided more physiologically relevant information about this system (Hayden et al., 2008).

NF-kappaB family of proteins (Perkins, 2007) is divided into two subfamilies, the NF-kappaB and the Rel subfamilies: the former contains three elements (p100 and p105, plus Relish in Drosophila) and the latter five (RelA, RelB and c-Rel, plus Dorsal and Dif in Drosophila). A common feature between these two subfamilies is the presence of a Rel homology domain (RHD) (Gilmore, 1990) that can serve both as a DNA binding and a dimerization domain. Proteins belonging to REL subfamily have a C terminus transactivation domain (TAD) that is responsible for the positive regulation of genes and that displays high variability allowing the regulation of many different genes (see further). NF-kappaB subfamily proteins usually are not transcription factors: however, they present at the C terminus ankyrin repeats that regulate the interaction with other proteins, therefore they can act as transcription factors when they form heterodimers with Rel subfamily proteins (Chen et al., 1999). In this last case the protein is subjected to limited proteolysis (p105 and p100 are respectively transformed into p50 and p52) ankyrin repeats are able to block the transcriptional activity of THDs (Hayden et al., 2004). Among these dimers the most relevant and most frequently studied is p50/ Rel-A, also known as p50/p65 dimer and very often referred as NF-kappaB itself. The activity of NF-kappaB as a transcription factor is regulated by Ik-B cytosolic proteins that form complexes with NF-kappaB: according to structural studies, IkB and NF-kappaB are in close contact burying nuclear localization signal of Rel, whereas the one of p50 is left exposed, ensuring a fine control on the shuttling of an inactive form of complex from the cytosol into the nucleus and vice versa, avoiding unwanted gene

expression. The complex is transiently activated by IkB kinase (IKK) that phosphorilates IkB (Hayden et al., 2004) and leads to its proteosomal degradation and to the shuttling of p50/p65 complex into the nucleus, and hence to gene expression through its interaction with 9-10 base pairs DNA sites (kB sites) (Gilmore, 2006). This NF-kappaB signal transduction pathway is called canonical pathway (Hayden et al., 2008); other mechanisms have been identified (*i.e.*, non canonical pathway fo T cells activation) or are under investigation (Perkins, 2006). The three mechanisms are reported schematically in Fig.1.4.

1.3.2 Natural drugs against NF-kappaB

Due to the long term tradition behind them and to the spread usage of them, natural drugs are compounds whose primary and side effects are very well known and documented. Beside a knowledge provided by popular medicine, more recently these molecules have been investigated adopting contemporary scientific standards in order to provide detailed information about their biochemistry and their molecular biology activity (Maroon et al, 2006).

In this sense in the last decade there has been an intense research activity that led to an increased interest towards common plant derived compunds like polyphenols (Feng, 2006). Polyphenols are the principal constituents of green tea, a dietary supplement whose preventive effects towards cardiovascular diseases and cancer has been widely recognized in the last years and attributed to its antioxidant properties. In particular, main efforts have focused on the study of Epigallocatechin-3-Gallate (EGCG), the most abundant catechin of green tea (Kundu et al., 2005) (Fig. 1.5a). Several recent studies demonstrated that the inactivation of NF-kappaB by EGCG was associated with inhibition of IKK activity (Chen et al., 2002), enhancement of phosphorylation dependent degradation of I B and subsequent increase in nuclear translocation of p65 protein (Wheeler et al., 2004). In this sense it has been also recently suggested that EGCG could partially inhibit in epithelial cells, through the interaction with the NF-kappaB pathways, the release of IL-8, a primary neutrophils attractant during the

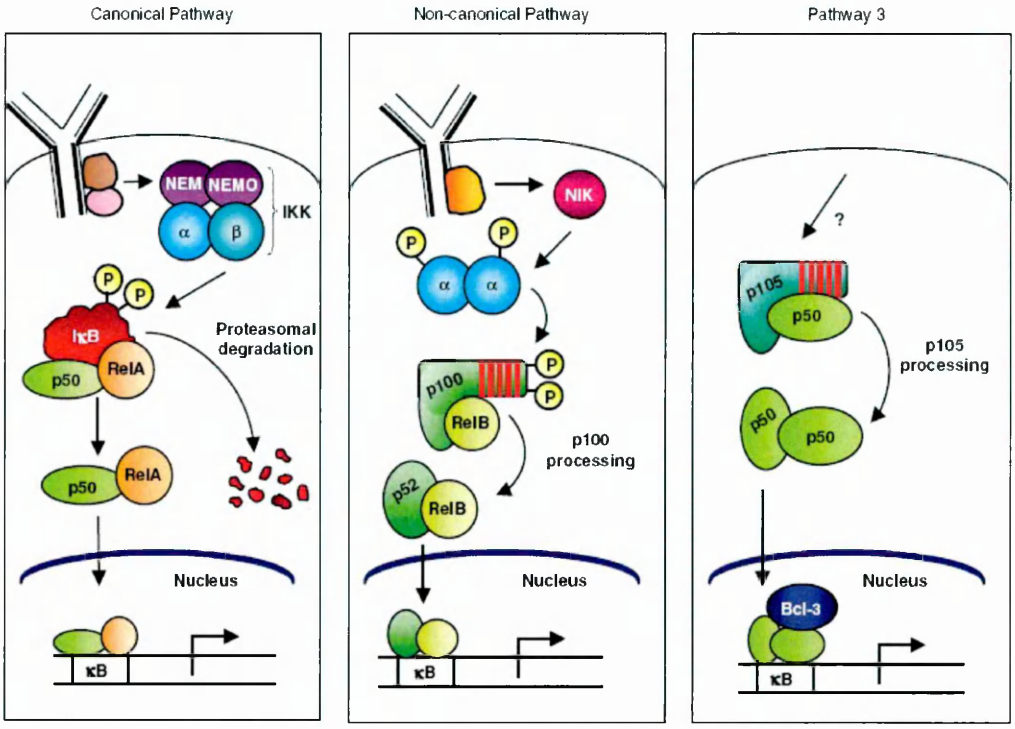


Fig. 1.4 NFKB pathways a. canonical pathways (see text); b. non canonical pathway: it islargelyconnected to the activation of p100/RelB complexes during B-and T-cell organ development, and it differs from the canonical pathway in that only certain receptor signals (e.g., B-cell activating factor or CD40) activate this pathway and because it proceeds through an IKK complex that contains two IKKa subunits ; c. new non canonical pathway: p50 (or p52) homodimers enter the nucleus, where they become transcriptional activators by virtue of interaction with the IkB-like co-activator Bcl-3 (or IkBz). How Pathway 3 is regulated is not known (from Gilmore, 2006).

inflammation process. Another study proposed also a T cells specific effect of ECGC in downregulating NF-kappaB: the drug is able to inhibit the catalytic activity of both 20S and 26S proteasome, as demonstrated *in vitro*, thus leading to an accumulation of intracellular IkB and to an inhibition of NF-kappaB activation . This has the consequence of blocking TNF alpha synthesis and to induce T cells cycle arrest, preventing any autoimmune response (Aktas et al., 2004).

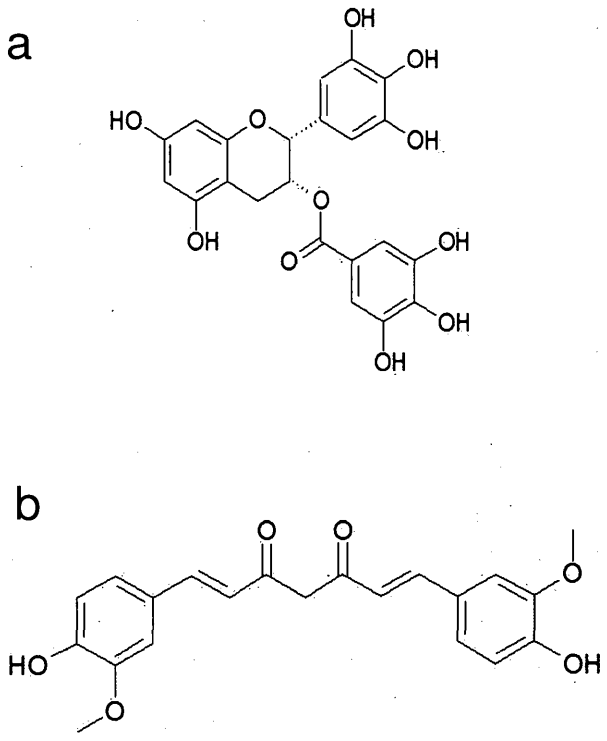


Fig. 1.5 a. EGCG molecule b. curcumin molecule.

Curcumin (difeluroyl methane, Fig.1.5b) is another compound that capitalized the experimental efforts of many pharmaceutical chemists and molecular biologists, as shown by recent publications in this field of research (Kundu et al., 2005). This molecule is the major component of many spices that are mainly members of the ginger family, among which curry is the most diffused one, especially after its introduction in western countries by means of the diffusion of Indian and Chinese food. The relevance of curcumin in biological processes has been already understood by ayurvedic medicine (I.e, the Indian traditional medicine) that makes large usage of curcumin for the topical treatment of skin wounds, herpes zoster, acne and parasitic skin infections (Strimpakos

et al., 2008). During 80s curcumin has been quite extensively biochemically studied and the antioxidant ability of the drug emerged rapidly from many studies (Strimpakos et al., 2008). Moreover, many evidences emerged from experiments performed on human platelets, colon epithelial cells and rat peritoneal neutrophils that indicated in curcumin an antiinflammatory drug that was able to interact with the pathway arachidonic acid through inhibition of COX-2 gene expression, thus acting at more nuclear level than synthetic COX-2 inhibitors (Plummer et al., 1999). More recently, however, it became more evident in the last five years that also the NF-kappaB pathway is influenced by the assumption of curcumin: in late 90s already a role for curcumin in NF-kappaB pathway has emerged in bone marrow stromal cells through TNF-a induced activation of NF-kappaB transcription (Kim et al., 2007), in agreement with the observation that in colon cells curcumin is able to inhibit the formation of the NIK/IKK signaling complex at the level of IKKa/IKKb dimer (Plummer et al., 1999), thus being able to prevent the overexpression of COX-2 especially in colon cancer. These latest discoveries allowed the approval of an experimental protocol for phase II chemopreventive trials for the treatements of patients with colon cancer with curcumin (Murakami et al., 2007).

1.3.4 Rediscovery of old drugs as NF-kappaB interactors

Natural drugs are very promising molecules for some diseases, as they can be easily integrated in normal diet and in some cases they can be used at high doses in clinical trials. However, clinical trials are very expensive, and only in the last twenty years natural substances are going to be reconsidered for their pharmaceutical relevant properties, due to the lack of real scientific information about their activity, as explained in the above section. In this sense, one of the latest discovery activity of pharmaceutical companies is the rediscovery of old drugs for diseases that have not been considered as their original targets. This procedure is very cheap because the clinical trials for these drugs have already been performed and for many of them side effects and posology are very well known. In some cases also long term side effects, *i.e.* the ones that have not been covered during clinical trials but only after huge usage by customers during the market life of the drug, are known and very well documented with a reliable statistical

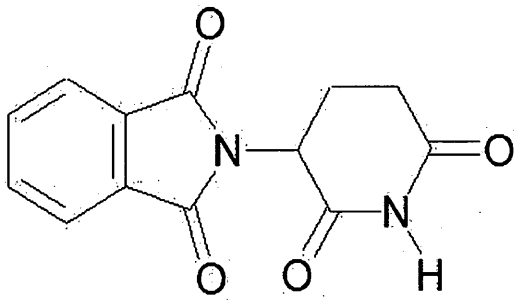
relevance. In this sense, in the last decade NF-kappaB seems to be the most relevant target of this branch of research for what concerns inflammation and antiinflammatory drugs, due to the dangerous side effects of COX-2 selective drugs and due to the aforementioned problems about inhibition of prostaglandines synthesis.

The most relevant and recent example of this rediscovered drugs is thalidomide. This compund ([+]-a-phthalimidoglutarimide, Fig.1.6a) was first synthesized almost 50 years ago as an antihistamine drug (Raje et al., 2002) and soon found to have sedative effects in animal studies. Due to lack of toxicity in animals at high doses, it was approved in 1957 as an over-the-counter sedative during pregnancy in almost 50 countries. However, this success did not last very long, because thalidomide was found to induce birth defects in humans; it was withdrawn from the market in 1965 (Meierhofer et al., 2001). About thirty years later thalidomide selective inhibition of TNF produced by stimulated human monocytes was demonstrated, and the central role of TNF in a wide variety of diseases placed this drug on a comeback trail (Ogata et al., 2003). In 1998, thalidomide was approved in the United States for the treatment of erythema nodosum leprosum, a complication associated with leprosy (Richardson et al., 2002).

In 1999 (Moreira et al., 1999) a decrease in endothelial cell proliferation in cultures treated with thalidomide or the thalidomide analogue cc-1069 was observed, The inhibition occurred in association with a marked decrease in the activity of the nuclear factor SP1 and a moderate inhibition of NF-kappaB activation in nuclear extracts of endothelial cells. These findings indicated that thalidomide and analogs could have anti-inflammatory, antiangiogenic and immunomodulatory properties. In the following years several papers were published that demonstrated the NF-kappaB inhibitory properties of thalidomide (Jin et al., 2002): this drug is able to inhibit the NF-kappaB associated release of IL-1β in colon cancer cells and, more in general, NF-kappaB activation is suppressed by thalidomide in different ways that depend on the proinflammatory stimulus used on cell cultures (*i.e.*, TNF and hydrogen peroxide stiumulation of NF-kappaB is inhibited by thalidomide, whereas no effect is observed if ceramide, phorbol esther or lipopolisaccharides are used). More recently, also in burn

injuries TNF levels appeared to be decreased by thalidomide (Eski et al., 2008). Due to the relevance of the NF-kappaB pathway on the multiple myeloma pathology in these last years, thalidomide has been experimentally investigated as a potential anti myeloma agent, demonstrating a potential antiapoptotic effect of this drug among others tested (Mitsiades et al., 2002).

a



b

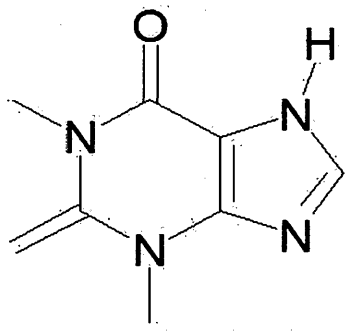


Fig. 1.6 a. Thalidomide molecule b. Theophilline molecule.

1.4 Looking for antiinflammatory targets: what's new

If all the above listed drugs could be a good choice for some alternatives to the usage of

more "standard" drugs in a near future and, in the case of natural drugs, a good way for preventing the degeneracy or even the appearance of inflammatory events, a question still remains open. Are there any other way to block inflammation finding genetic or biochemical pathways alternative to the ones historically targeted for anti-inflammatory purposes? In a more practical way, could it be any possibility of finding a novel protein target in order to elaborate a new and robust anti-inflammatory strategy? In this sense, very recently three proteins have been identified to be good candidate for modulating inflammatory response in the organism: histone deacetylases (HDACs), Interleukin 33 (IL33) and extracellular High Mobility Group Box 1 (HMGB1) protein, which is the object of the present work.

1.4.1 Theophyllin and HDACs

At nuclear level and within the process of chromatin remodeling, acetylation of core histones is a mechanism associated with activation and transcription of various genes and is regulated by coactivator molecules (HAC) that have intrinsic histone acetyltransferase activity (Urnov et al., 2001). Also some genes that express proteins involved in inflammation are regulated by acetylation: proinflammatory transcription factors such as NF-kappaB bind to coactivator molecules and activate this enzyme. Histone acetylation is reversed by HDACs: decreased HDAC activity also results in increased NF-kappaB mediated inflammatory gene expression without affecting NF-kappaB DNA binding (Ichiyama et al., 2001). This process is deeply studied since 20 years as a basic chromatin remodeling process, but its role in inflammation in inflammation only emerged in the last years. However, no drugs seemed to be involved in any interaction with HAC/HDAC pathway.

Theophylline (Fig.1.6b) is one of the most widely used drugs in the treatment of COPD (chronic obstructive pulmonary disease) since 1940s, and recently published guidelines for COPD recommend theophylline as one of the central drugs in the management of symptomatic COPD, despite the preference for inhaled β-agonists especially in the richest countries for some unavoidable side effects of theophylline. The mode of action of this drug remained unexplored and unclear for many years. A partial

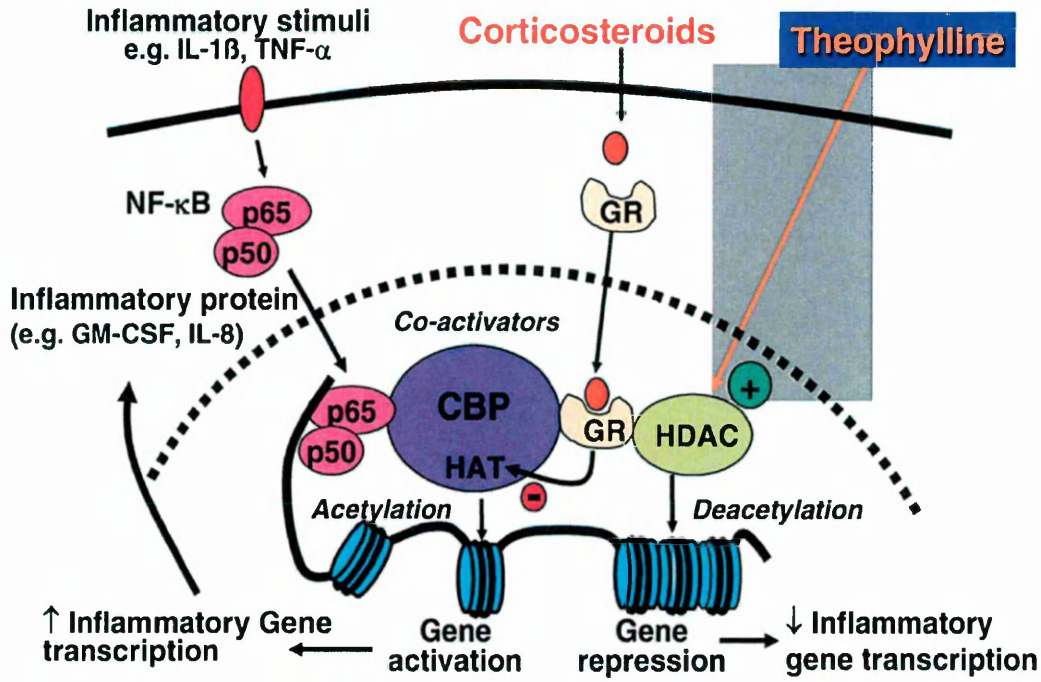


Fig.1.7 Theophylline directly activates histone deacetylases (HDACs), which deacetylate core histones that have been acetylated by the histone acetyltransferase (HAT) activity of coactivators such as CBP (cAMP-responsive element-binding [CREB] protein). This results in suppression of inflammatory genes and proteins, such as granulocyte-macrophage colony- stimulating factor (GM-CSF) and interleukin (IL)-8, that have been switched on by proinflammatory transcription factors, such as NF-B. Corticosteroids also activate HDACs but through a different mechanism, resulting in the recruitment of HDACs to the activated transcriptional complex via activation of the glucocorticoid receptors (GRs), which function as a molecular bridge. This predicts that theophylline and corticosteroids may have a synergistic effect in repressing inflammatory gene expression (from Barnes, 2005).

antiinflammatory effect was already carachterized starting from the end of 90s and related to the NF-kappaB mediated expression of IL-8, but this effect was only observed in cells at low concentration of theophylline (Tomita et al., 1999).

Very recently, some publications revealed that in COPD NF-kappaB activation is increased in macrophages in induced sputum, peripheral lung and airway epithelial cells; moreover, there is a marked reduction in HDAC activity in COPD alveolar macrophages, bronchi, and peripheral lung (Di Stefano et al., 2002) and decreased HDAC activity also results in increased NF-kappaB mediated inflammatory gene expression without affecting NF-kappaB DNA binding (Zhong et al., 2002). Corticosteroids suppress the expression of inflammatory genes by binding to and

activating glucocorticoid receptors, which recruit HDAC2 to the transcription complex of inflammatory genes that is activated, thereby reversing histone acetylation and silencing genes that have been activated by inflammatory stimuli (Ito et al., 2000). These mechanisms appear to account for most of the antiinflammatory effects of corticosteroids in asthma: however, in COPD the marked reduction in HDAC activity and expression may account for the characteristic resistance to the antiinflammatory effects of steroids seen in this disease (Barnes et al., 2004).

Theophylline activates HDAC activity and therefore suppresses the expression of inflammatory genes (Ito et al., 2002a) (Fig. 1.7) It is not yet certain whether HDACs are the direct target of Theophylline, because several other nuclear proteins are coprecipitated in these inflammatory gene complexes. Theophylline activates different subtypes of HDAC, with equal activation of HDAC1 and HDAC2 (Ito et al., 2002b).
The mechanism whereby low concentrations of theophylline activate HDAC is not yet known. The effect of theophylline on HDAC activity appears to be potentiated under conditions of oxidative stress. In the sense of a rediscovery of old or traditional drugs, the aforementioned novel possible usage of Theophylline is illuminating for two reasons: 1. a new mechanism has been identified that switches gene regulation at molecular level but with a different mechanism with respect to the most studied process that involves NF-kappaB; 2. an old drug that has been (and still it is) used for a purpose (i.e., bronchodilation) in this new context finds a new role as an antiinflammatory drug. However, despite the promising features, major problems in adopting theophilline as a new antinflammatory drug are the side effects of Theophylline has. Moreover, histone acetylation/deacetylation is a very delicate biochemical process, hence interacting directly with it could be very dangerous for the cell cycle and gene expression. Hence, if on one hand this is a possible route for saving money and have a ready to go drug, some severe questions are still open, and new targets are worth to be investigated and many other data must be collected in the near future.

1.4.2 IL33 boosting

In 2005 a paper published by the journal Immunity reported evidences that IL-33, an

interleukin belonging to the family of IL-1, is able to bind to IL-1 receptor ST2 on the surface of T Helper cells and hence to induce the release of proinflammatory cytokines from stimulated cells (Schmitz et al., 2005). In mid 90s two different peptidic ligands were discovered (Kumar et al., 1995; Gayle et al., 1996), but they did not belong to the interleukin family and did not display any biological, activity, as they were not able to activate NF-kappaB pathway. IL-33, instead, was demonstrated, both *in vitro* and *in vivo*, to both directly bind ST2 and to activate NF-kappaB pathway through phosphorylation and MAP kinases in mast cells. This is reflected in the production of cytokines by T helper cells, and anatomical changes were observed in the lung and in the digestive tract in mice. In the following years, there have been many efforts in the direction of providing a more detailed description of the IL33-ST2 interaction and of its pathological and physiological relevance because of the influence of this complex on inflammatory processes in cells. In this sense, also the chemoattractant activity of IL-33 towards T helper cells has been demonstrated and evidences have been found of the IL-33 stimulation of cytokines release also by mast cells. IL-33 is involved at least in three different pathologically relevant mechanisms: IL33-ST2 interaction has bee detected *in vivo* in animal models (murine thymoma cell line) of airway inflammation and the same release of cytokines previously described has been observed; b. in cardiac heart cells IL-33/ST2 signaling is biomechanically (due to the mechanical contact between cardiac fibroblasts and cardiomyocytes) activated and it controls the cardiomyocyte hypertrophy and cardiac fibrosis; c. IL-33 has a role in atherosclerosis prevention because it is able to induce in high fat diet mice the production of antibodies against IL-5 and oxidized LDL, hence preventing the formation of typical atherosclerotic plaques in vessels.

From a molecular point of view, IL-33 is a dual function endothelium derived protein that may function as both a proinflammatory cytokine and an intracellular nuclear factor with transcriptional regulatory properties: IL-33 is identical to NF-HEV, a nuclear factor preferentially expressed in high endothelial venules (HEV), it is a heterochromatin-associated nuclear factor in HEV endothelial cells *in vivo* and it possess transcriptional repressor properties, associated with the homeodomain-like helix–turn–

helix motif. All these features, taken together, demonstrate that IL-33 and its interactions with its receptor ST2 could become central in the development of an antiinflammatory drug in the next future. However, up to date, no information about IL-33 structure are available nor an atomic level description of interactions between IL-33 and ST-2, which are the necessary prerequisites for drug discovery and/or design.

1.5 HMGB1: from nucleus to outside the cells triggering inflammation

1.5.1 HMGB1 and the nucleus: a story of long ago

Like IL-33, also HMGB1 has the dual role of architectural chromatin binding protein and of cytokine, making HMGB1 a possible interesting target for drug discovery.

HMGB1 was initially discovered almost 30 years ago and in early stages of research it was named amphoterin. It was readily identified as an abundant (up to a milion molecules per cell, almost 10% of the population of nuclear histones) nuclear protein (expressed in mammals, yeast, bacteria and plants) that is important for the regulation of gene transcription and for chromatin remodeling and architecture (Rauvala, 2007). The human gene of HMGB1 is located on chromosome 13q12 and encodes 216 residues divided into two domains named Box A and Box B, approximately of 70 residues each, and a 30 residues acidic tail that, in the protein in its resting state, interacts with basic stretches in Box A and B, shielding them from other interactions that might occur before HMGB1 binds DNA (Fig.1.8). The knockout mouse of HMGB1 is lethal, although HMGB2 and HMGB3 molecules can provide redundancy during embryogenesis (Calogero et al., 1999) . However, HMGB1 binds rather weakly to the B-form variety of DNA, and has practically no sequence specificity. Instead, it binds with high affinity to unusual DNA structures like 4-way junctions and DNA bulges; this suggested that HMGB1 might actually function by distorting linear DNA into a bent conformation. The considerable energy required for DNA bending comes from the extensive contacts of HMG boxes within the minor groove of DNA. LEF-1 and SRY - transcription factors that incorporate an HMG Boxes - also bend DNA as part of their activity; the difference in sequence specificity between LEF/SRY and HMGBs has now been traced to specific amino acid residues (Lotze et al., 2005) .

Obviously, HMGB1 will interact in the nucleus with distorted DNA structures, such as those created by the incorporation of nucleotide analogues, by covalent modification or by adducts formed upon UV irradiation (Agresti et al., 2003) . HMGB1 shields cisplatin-damaged DNA from repair processes (He et al., 2000). Important as they are, these functions do not justify the abundance of HMGB1; whether HMGB1 is involved in homologous recombination is still undecided, although it can promote the non homologous integration of exogenous DNA into the genome (Ueda et al., 2002) . HMGB1 causes DNA bending and facilitates the binding of several regulatory protein complexes to DNA, particularly members of the nuclear hormone-receptor family, V(D)J recombinases and p53–p73 transcriptional complexes. HMGB1 overexpression, which is observed in many tumor cells, accelerates cell-cycle progression, presumably through its ability to promote access to some (and not other) transcriptional complexes. Interestingly, in some tumor cells, HMGB1 can be polyADP-ribosylated, a common phenotype change in proteins present in cells with damaged DNA associated with the invasiveness of tumors. HMGB1, similar to the functionally related bacteriophage proteins IHF and HU, stabilizes nucleoprotein complexes simply by its capacity to maintain a bend in DNA. Indeed, HMGB1 and members of the HMG domain protein family have unusual relationships with viruses: in some cases, they inhibit viral integration into the genome by binding viral DNA; and in other cases, they are targets for early viral inactivation (Lotze et al., 2005).

HMGB1 is very mobile within the nucleus, being able to traverse the entire compartement in less than 2 seconds due to a very transient association to chromatin associated to a jumping mode of binding of HMGB1 to different nucleosomes.

1.5.2 From nucleus to outside

HMGB1 does not contain any signaling sequence, hence it can not pass through endoplasmic reticulum or Golgi apparatus. However, secretion (or active release) of HMGB1 has been observed in many types of cells like macrophages, monocytes, neutrophils, platelets, dendritic cells and NK cells stimulated with LPS or with proinflammatory cytokines like TNF, IL-1β or IFNγ. This is a rare but not uncommon

event, since also other proteins (*e.g.*, IL-1b and FGF) are released without displaying a consensus sequence (Yang et al., 2005). Being independent from common secretory pathway, active release of HMGB1 depends on the relocalization from the nucleus to special organelles, the secretory lysosomes (Bonaldi et al., 2003). Moreover, it has been demonstrated that HMGB1 is relocated from nucleus through cytosol after acetylation induced by LPS or, in resting macrophages, by acetylases hyperacetylation (Bonaldi et al., 2003). In addition to active release, monocytes and macrophages are able to

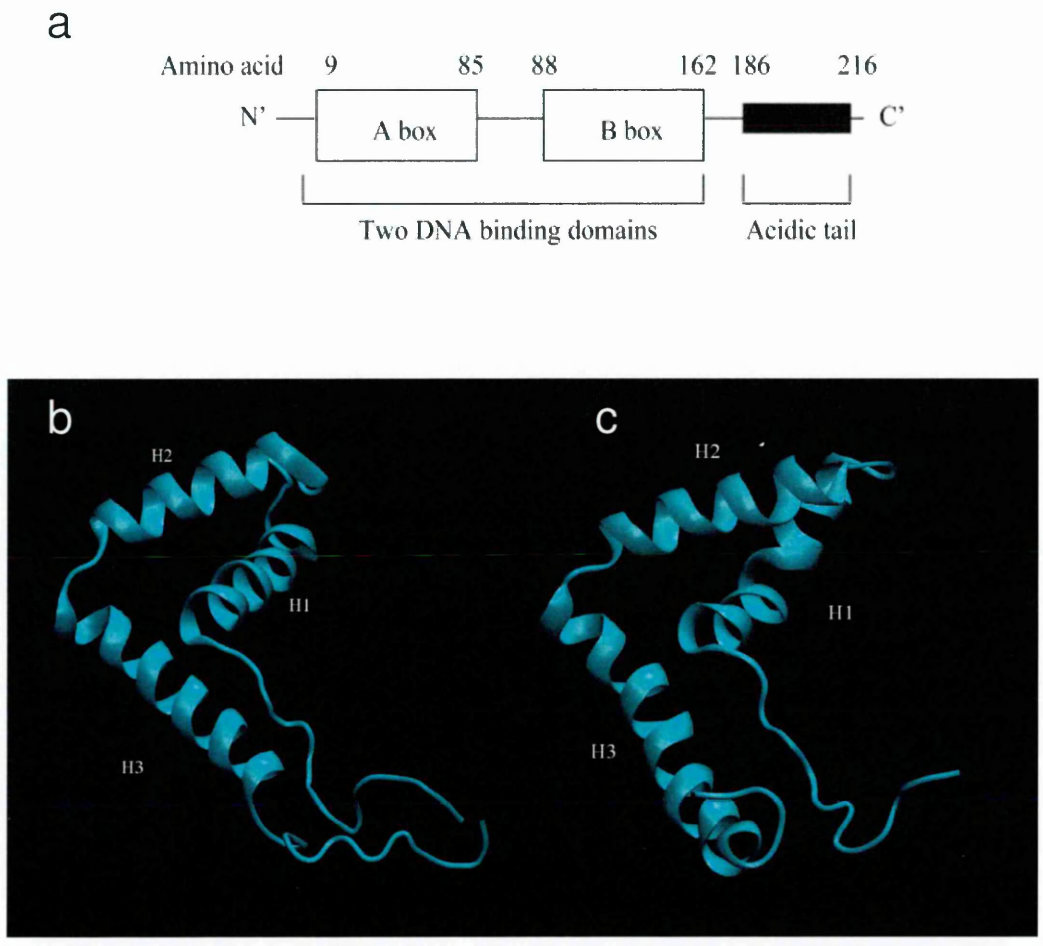


Fig. 1.8 HMGB1 structure and organization a. schematic representation of HMGB1 (from Yang et al. 2005); b. structure of the Box A of HMGB1; c. structure of the Box B of HMGB1.

passively secrete HMGB1 during necrosis. Cells that die in a programmed way (apoptosis) retain HMGB1 bound to chromatin remnants even after eventual cell lysis

(Wang et al., 1999) and its motion is blocked: this behavior is characteristic of HMGB1 and it has never been observed for other nuclear proteins. During apoptosis, HMGB1 is not post-translationally modified (Fig.1.9a).

In contrast, when cells die in a nonprogrammed way (necrosis), they release HMGB1 by simple diffusion due to cytoplasmic membrane disruption. HMGB1 deficient necrotic cells display a reduced inflammatory reponse if compared to wild type cells. This differential behavior between necrotic and apoptotic cells makes HMGB1 the primary signal of tissue damage at least at injury site; extracellular HMGB1 promotes local and systemic responses in the organism, including inflammation and the activation of innate and adaptive immunity (Scaffidi et al., 2002)(Fig.1.9b).

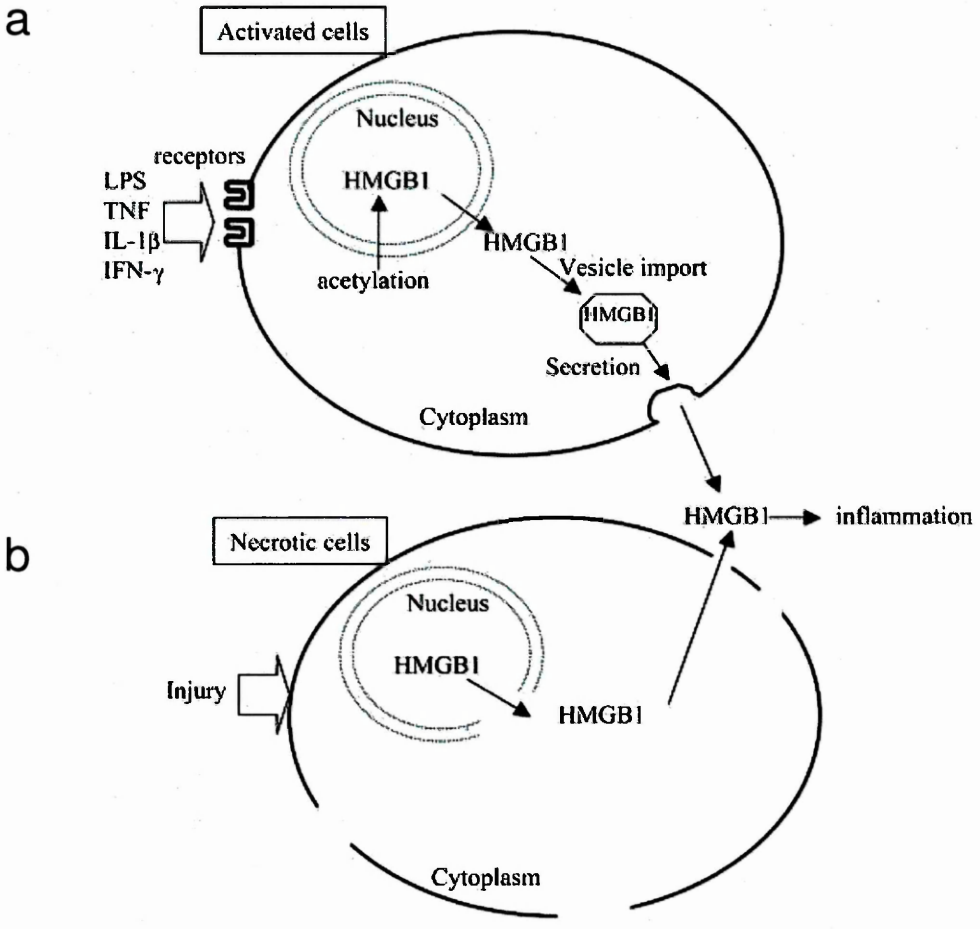


Fig. 1.9 HMGB1 release during cell death : a. active release; b. passive release (from Yang et al., 2005).

1.5.3 Inflammatory response in vitro and in vivo

Extracellular HMGB1 activates a large number of different physiological responses in different cell types, and can be considered a cytokine (Muller et al., 2004). *in vitro* there are experimental evidences that macrophages, monocytes, endothelial cells, neutrophils, epithelial cells and dendritic cells can have a cytokine release activity related to HMGB1 during the response to inflammation stimuli. In particular, as inflammation is primarily mediated by macrophages, monocytes and endothelial cells, the most numerous publications in the field are about these cellular systems: due to the stimulation with HMGB1, a huge ensemble of cytokines like TNF, IL-1β, IL-1α, IL-6, IL-8, MIP-1α and MIP-1β are released by the first two types of cells, whereas endothelial cells release mainly ICAM-1, VCAM-1 and RAGE (see further), but also TNF and IL-8 like macrophages and monocytes. Interestingly, in cultures of umbilical endothelial human cells, the proinflammatory activity of the Box B of HMGB1 was demonstrated. Moreover, HMGB1 is able to interact destruptively with a endothelial monolayer, this suggesting a role of the protein in promoting inflammatory stimuli in the injury site of a tissue (Loetze et al., 2001)

In vivo HMGB1 inflammatory effects have been experimentally detected in brain, lung and joints. Also in inestinal tract (Sappington et al., 2002) and in heart (Germani et al., 2007) HMGB1 has a role in regulating inflammatory response, but in these last cases this role is still not clearly documented. In particular, in lung HMGB1 levels are very high in patients and mice with acute lung injury: in animal models the injection of HMGB1 lead to an increase of levels of TNF, IL1b and MIP-2 and to lung edema, whereas antibodies against HMGB1 are able to decrease both lung edema and neutrophil levels. Joints related HMGB1 studies are very important due to the intimate correlation with reumatoid arthritis: high HMGB1 levels have been found in the sinovial fluid of reumatoid arthritis patients, whereas in adjuvant arthritis animal models HMGB1 is expressed in all cellular compartements compared to control animals where only nuclear HMGB1 is detected (Taniguchi et al., 2003).

1.5.4 HMGB1 receptor(s)

If the nuclear life of HMGB1 has been extensively characterized and the effect of HMGB1 on different types of cells has been fully investigated and it is well documented, many questions about the receptor of HMGB1 are still open. Historically, the first identified HMGB1 receptor was the receptor for advanced glycation endproduct (RAGE), a transmembrane protein belonging to the family of immunoglobulins and involved in many different receptor-ligand interactions (Bierhaus et al., 2006). RAGE is expressed in endothelial cells, in macrophages/monocytes and in vascular smooth cells, and its interactions with HMGB1 has been reported in tumour cells and in macrophages. In general, from a signaling point of view, RAGE is able to signal for the activation of two proteins, CDC42 and Rac, that regulate cell motility and neurite outgrowth (Huttunen et al., 2002). Moreover, RAGE is also able to activate a cascade of signals through the phosphorylation of many MAP kinases (p38, p42/44, ERK) and finally through the activation of the NF-kappaB pathway (Huttunen et al., 2002). Residues belonging to the C terminal part of HMGB1 Box B have been indicated as possible responsible for the interaction with RAGE. However, all these aforementioned evidences are not sufficient in order to explain all the biological effects of extracellular HMGB1. Moreover, there are serious doubts that RAGE is, at least, the only receptor of HMGB1 mainly because there are some experimental indications that show that HMGB1-mediated effects are only partially reduced by blocking RAGE (*i.e.*, by antibodies against the receptor) (Li et al., 2003). More recently some results evidenced that Toll like receptors (TLR) of type 2 and 4 are able to interact with HMGB1 leading to the activation of the NF-kappaB pathway. However, the implications of these interactions are still under investigation, like the hypothesis of a synergistic effect of RAGE and TLRs in binding extracellular HMGB1 (Yu et al., 2004).

1.5.5 Pathological relevance of HMGB1

HMGB1's importance in pathological aspects of inflammation has been already demonstrated in systemic inflammation diseases like sepsis, reumatoid arthritis and atherosclerosis.

The systemic release of HMGB1 in murine models of sepsis (Wang et al., 1999) has been largely documented, and also in human patients HMGB1 levels are much higher than in healthy subjects. Interestingly the kinetics of HMGB1 release is delayed with respect to other cytokines, like the HMGB1 induced release of TNF in macrophages, hence indicating that HMGB1 does not act as an early acting cytokine, but its effect comes late after the injury.

In reumatoid arthritis, HMGB1 is found at high levels in junctions and in particular in sinovial fluids: injection of HMGB1 in joints of healty mice induce cytokines release by macrophages, including TNF, IL-1b and IL-6. These data suggest a role of HMGB1 in the cartilage damaging process typical of arthritis, inducing some groups in the recent years to investigate HMGB1 as a marker of inflammation at least in this pathology (Taniguchi et al., 2003).

In the last five years also a role for HMGB1 has been also found in atherosclerosis (Porto et al., 2006), due to the release of HMGB1 by smooth muscle cells (SMC) in human atheroscleotic plaques. SMCs proliferate, migrate, and actively secrete more HMGB1 than other cells in response to HMGB1 itself and, partially, in response to cholesterol loading.

1.5.6 Beneficial effects of HMGB1

Not only pathological detrimental effects of HMGB1 are known, but also beneficial ones. Extracellular HMGB1 induce both migration and proliferation of vessel-associated stem cells (mesoangioblasts) both *in vitro* and *in vivo* (Palumbo et al., 2005). It was demonstrated that alpha-sarcoglycan null dystrophic muscle cells contain elevated levels of HMGB1, and mesoangioblasts were able to migrate into dystrophic muscles even if their RAGE receptors were disabled. This implies that the HMGB1-RAGE interaction is sufficient, but not necessary, for mesoangioblast homing, hence HMGB1 can in principle be used in order to induce muscle tissue regeneration by simple injection.

Mouse models demonstrated that HMGB1 can cover an important role in left-ventricular function and myocyte regeneration after infarction (Limana et al., 2005): administration of HMGB1 in the left ventricle of an operated animal resulted in the

formation of new myocytes within the infarcted portion of the wall. The regenerative process involved the proliferation and differentiation of endogenous cardiac c-kit+ progenitor cells. Echocardiographic and hemodynamic parameters at 1, 2, and 4 weeks demonstrated a significant recovery of cardiac performance in HMGB1-treated mice.

1.5.7 Anti-HMGB1 therapy

Up to date, a real anti-HMGB1 drug does not exist, and a treatment of HMGB1 involving disases or pathology is still missing. However, many steps in this direction have been made.

Antibodies against HMGB1 can reverse sepsis caused by peritonitis in mice (Kokkola et al., 2003; Yang et al., 2004); the effect of HMGB1 is dose dependent and it is effective only after TNF clearance during inflammation. Surprisingly, such a similar protective effect was observed also for the HMGB1 Box A, that thus acts as an anatagonist of HMGB1. These two "protein derived drugs" have also some beneficial effects on collagen induced arthritic pathologies in rodents, thus suggesting and anti-HMGB1 strategy for the treatement of such a disease. Moreover, ethylpiruvate, a non toxic food additive, has been also tested as anti-inflammatory agent: it is able to inhibit the LPS induced HMGB1 release and the activation of the NF-kappaB and p38 MAPK pathways in macrophages (Ulloa et al., 2002).

1.5.8 Triterpenic drugs and interactions with HMGB1

Triterpenic drugs are the biochemical product of the tail to tail condensation of two molecules of pharnesyl pyrophosphate. They are traditionally named "saponins" because of their ability to decrease the surface tension of solutions (Goegelein et al., 1984) and to produce bubbles in them like soap: for this reason saponins are often used in chemical industries as surface pressure reducer. The most relevant biological property of saponins is the emolysis of eritrocytes (Baumann et al., 2000). In humans saponins are only poorly absorbed by intestinal walls and this makes them not toxic if orally administered (Samuelsson, 1999). Chemically, triterpenic saponins are characterized by the presence of a triterpenic scaffold (aglicone) connected through an ether bond to oligosaccaride

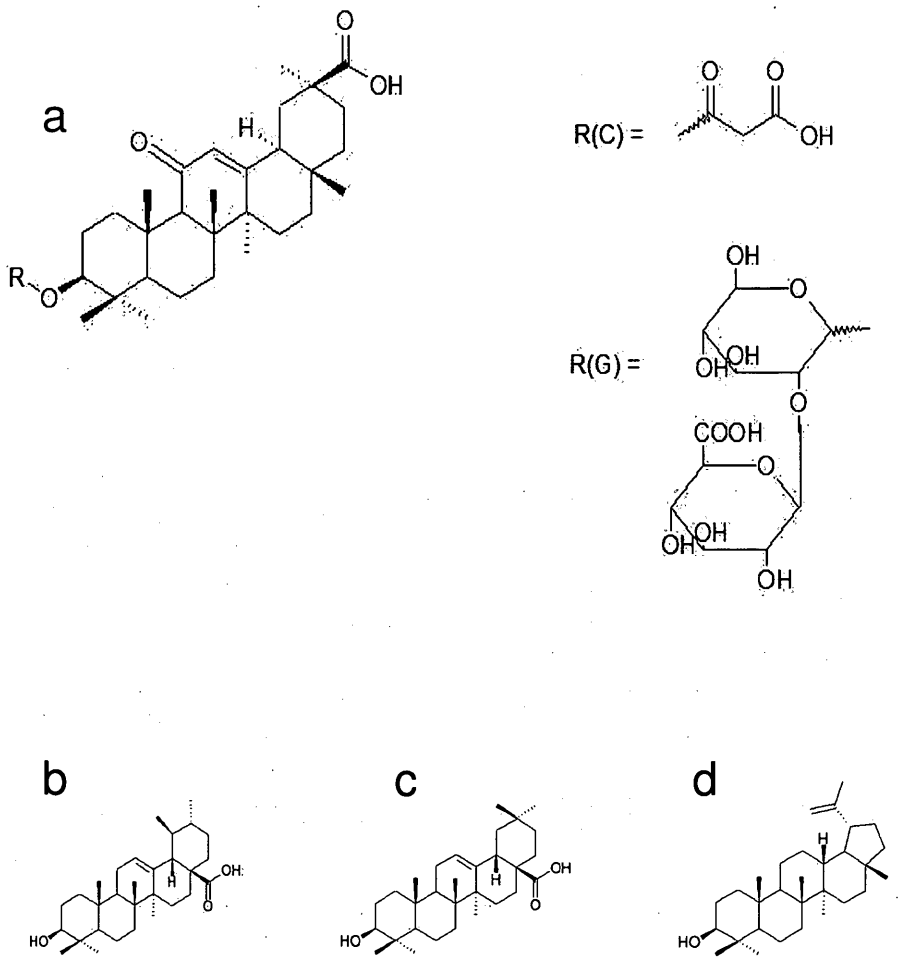


Fig. 1.10 a. Structure of carbenoxolone [R(C)] and of glycyrrhizin [R(G)]; b. ursolic acid; c. oleanoic acid; d. lupeol.

chain composed of up to 12 sugar units (mono saccharides or uronic acids). If the triterpenic scaffold is substituted by steroidal scaffolds the whole class of compound is referred as "steroidal saponins" and it is the major natural source of hormones for pharmaceutical usage. One of the most important triterpenic natural drug is glycyrrhizin (Fig.1.10a), whereas oleic acid, ursolic acid and lupeol are the three major sources of triterpenic scaffold for triterpenic saponins (Fig. 1.10b,c and d).

Glycyrrhizin (GL) is the name of the sodium or potassium salt of the

glychirrizhic acid, whose triterpenic scaffold is called glychiritinic (or glychirretic) acid (Fig.1.10a). The major sources of GL are the licorice root and the licorice extract, that contain up to 90% of the active principle and that can be obtained by 3 or 4 years old *Glychirriza glabra*, of the *Fabaceae* family, a 1-2 m tall plant that grows in the Mediterranean Sea area, in the South of Russia, in Iraq and in Iran. Licorice roots are grey-brown and are longitudinally striped; licorice extract is a viscous mass that is usually conserved in rods (Samuelsson, 1999). Other triterpens like b-amyrin, stigmasterol, onocerin and sitosterol are present in minor quantities in the licorice roots/extracts together with 30-40 different types of flavonoids, a small amount of cumarin and all the precursor of GL at small percentage (see Fig.1.11) (Samuelsson, 1999).

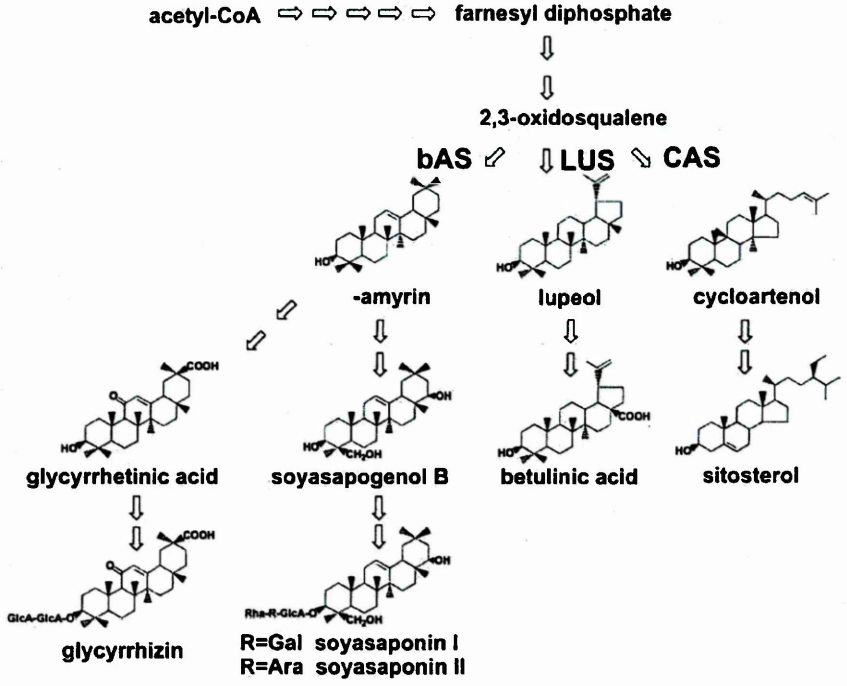


Fig. 1.11 Glycyrrhizin biosynthesis in plants (from Hayashi et al., 2004).

GL is widely used in food industry because of a taste 50 times sweeter than saccharides. Pharmacologically it has a relevant anti inflammatory and anti ulcer activity (see further), a still controversial expectorant activity and an anti diuretic activity, despite the fact that a prolonged usage can generate edemas. More in general side effects of GL are very similar to the ones of corticosteroids, *i.e.* hypertension, hypopotaxemy, lowered glucose tollerance (Samuelsson, 1999, Matsui et al. 2004).

Other sources of pharmacologically relevant triterpenic saponins are the roots of *Quillaja saponaria*, of *Poligala senega* and of *Panax ginseng*. The active principle of the first two plants are mainly used as expectorant; ginsenosids are derivative of 20(S) - protopanaxadiol and of 20(S) – protopanaxtriol and are adaptogen drugs whose pharmacological activity towards platelets aggreagation, inflammation, CNS stimulation or depression has been documented in the last decade (Samuelsson, 1999).

For the past few decades, the GL preparation Stronger-Neo Minophagen C (SNMC) has extensively been used to treat chronic hepatitis in Japan (Arase et al. 1997) , and the safety and efficacy of SNMC have established. SNMC is now under clinical trial in Europe in patients with chronic hepatitis C (van Rossum et al., 2001) . So far, GL has been demonstrated to have a variety of activities, such as the production of several cytokines, namely IFN-g, IL-2, IL-12, and IL-10, and the augmentation of NK cell/NKT cell activity (Matsui et al., 2004) . Also, GL administration induces anti-inflammatory, cytoprotective effects *in vivo* and anti-viral effects *in vitro* (Shibata et al., 2000). In early 2003, severe acute respiratory syndrome (SARS) raged in several countries. Surprisingly, it was promptly reported that GL was effective against SARS-associated coronavirus (Cinatl et al., 2003). A subsequent paper reported that an elevation of plasma inflammatory cytokines and chemokines was observed in SARS patients (Wong et al., 2004) and more recently it was experimentally demonstrated that GL and related compounds are able to down-regulate production of inflammatory chemokines IL-8 and eotaxin 1 in human lung fibroblasts (Sato et al., 2001).

Carbenoxolone (CBNX), a synthetic derivative of GL, is a licensed drug used for oesophageal ulceration and inflammation. Recently it has also been investigated for

nootropic effects (Sandeep et al., 2004): moving from observations about negative effects on cognition and memory generated by long-term exposure to glucocorticoids, it was possible to monitor a CBNX induced decrease of the amount of active glucocortocoid in the brain due to the 11Beta-hydroxysteroid dehydrogenase type 1 inhibition. CBNX side effects are the same of GL.

In 2001 (Sakamoto et al., 2001) it was demonstrated *in vitro* that there is a physiological correlation between GL and high mobility group proteins 1 and 2 (HMG1 and HMGB2) and the inhibitory effect of GL on their phosphorylation by three protein kinases (CK-I, CK-II and PKC). It was found that HMG1/2 have a high affinity with a GL-affinity column and that GL induces a small conformational change in HMG1/2 (revealed by a change in the CD spectrum between the free and the bound state of the protein). Phosphorylation of HMG1/2 by two protein kinases (CK-I and PKC) was completely inhibited by a glycyrrhetinic acid derivative (oGA) at one-tenth the concentration of GL. Also, the DNA-binding abilities of HMG1/2 were reduced by GL in a dose-dependent manner. These results suggest that the binding of GL to HMG1/2 results in the inhibition of their physiological activities (DNA-binding ability and phosphorylation by PKC or CK-I) *in vitro*.

More recently, G.Sitia and collaborators (Sitia et al., 2007) demonstrated *in vivo* the effect of GL on HMGB1 in Hepathitis B Virus (HBV) mice models. The authors found that HMGB1 is not involved in the migratory and antiviral activity of virus-specific , but it is necessary for the intrahepatic recruitment of antigen non-specific cells that amplify the cytotoxic T lymphocytes (CTL) dependent liver damage, suggesting that a similar role of HMGB1 could be played in the pathogenesis of viral hepatitis in man. Since the treatment of mice with Box A-A or GL is associated with maintenance of CTL-dependent antiviral effects but diminished tissue damage, the design of new and more potent inhibitors of HMGB1 may be useful for the treatment of chronic HBV infection. Along these lines these results shed new light into the mechanism whereby *GL* may carry out its beneficial effect.

1.6 Preliminary results

1.6.1 Glycyrrhizin inhibits the chemoattractant Activity of HMGB1

HMGB1 has chemoattractant activity on endothelial, smooth muscle, and vessel-associated stem cells (Palumbo et al., 2004). In addition, Marco E. Bianchi and coworkers (personal communication) demonstrated that it is also a powerful chemoattractant for the widely available mouse 3T3 fibroblasts (Fig. 1.12). GL interfered with HMGB1-elicited cell migration of 3T3 fibroblasts in a dose-dependent manner ($p < 0.001$ in ANOVA two-way test). The number of 3T3 cells migrating in

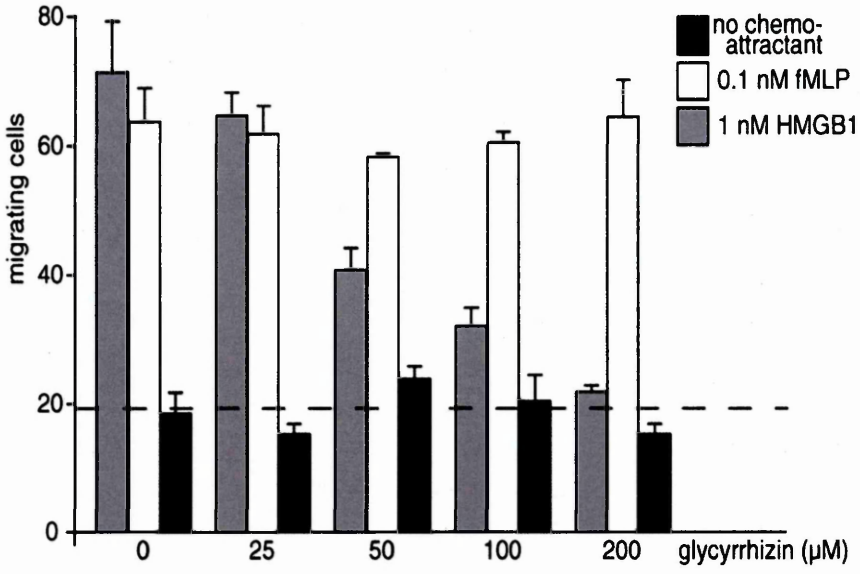


Fig.1.12 3T3 fibroblasts were subjected to chemotaxis assays in Boyden chambers where 0.1 nM fMLP, 1 nM HMGB1, or no chemoattractant was added in the lower chamber, together with the indicated concentrations of glycyrrhizin. The data represent the average ± SD of three replicates; the experiments were replicated at least three times. The inhibitory effects of glycyrrhizin on HMGB1-induced cell migration were highly significant ($p < 0.001$ in ANOVA analysis), whereas they were not statistically significant on fMLP-induced cell migration.

response to 1 nM HMGB1 was reduced by about half at a concentration of 50 mM GL (IC50 = 49 ± 1 mM in a sigmoidal log dose response curve, r2 = 0.93). GL did not affect cell migration elicited by fMLP, a well-known chemoattractant, even at a concentration of 200 mM. This shows that GL does not affect the general mobility of fibroblasts, and suggests that its effect is specific on HMGB1-elicited chemotaxis. Moreover, GL had no

statistically significant effect on the mobility of fibroblasts in the absence of chemoattractant.

HMGB1 promotes the *in vitro* proliferation of vessel associated stem cells (mesoangioblasts) in the absence of fetal calf serum (Palumbo et al., 2004). When added to mesoangioblasts in culture, increasing concentrations of GL inhibited the mitogenic

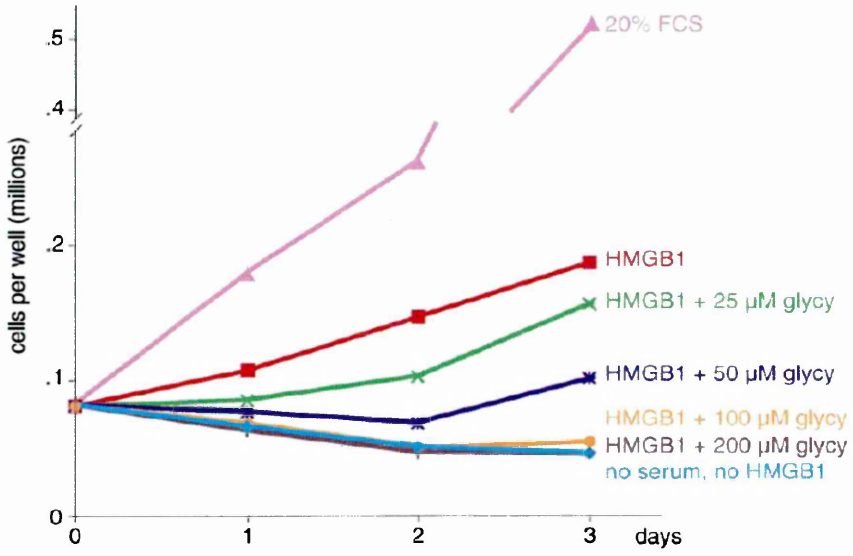


Fig.1.13 D16 mesoangioblasts were grown in RPMI medium containing no additions, 1 nM HMGB1 with the indicated concentration of glycyrrhizin, or 20% fetal calf serum (FCS). Each point represents the mean of three replicates. Error bars are omitted to avoid clutter; the standard deviation was between 2% and 15% of the average. The experiment was repeated three times. The inhibitory activity of glycyrrhizin HMGB1-induced stem-cell proliferation is highly significant ($p < 0.001$ in ANOVA analysis).

activity of HMGB1 in a dose-response manner; in the presence of 200 mM GL no mitogenic activity of HMGB1 was noted, and the cell counts were indistinguishable from those of cultures where no serum and no HMGB1 were present (Fig. 1.13).

1.6.2 Glycyrrhizin does not interfere with the nuclear function of HMGB1

In order to test if GL might interfere with the DNA-binding function of HMGB1, fluorescence recovery after photobleaching (FRAP) experiments in order to monitor whether GL would increase the mobility of intranuclear HMGB1, and reduce its residence time on chromatin (Fig. 1.14). In the presence of 100 mM of GL, the mobility

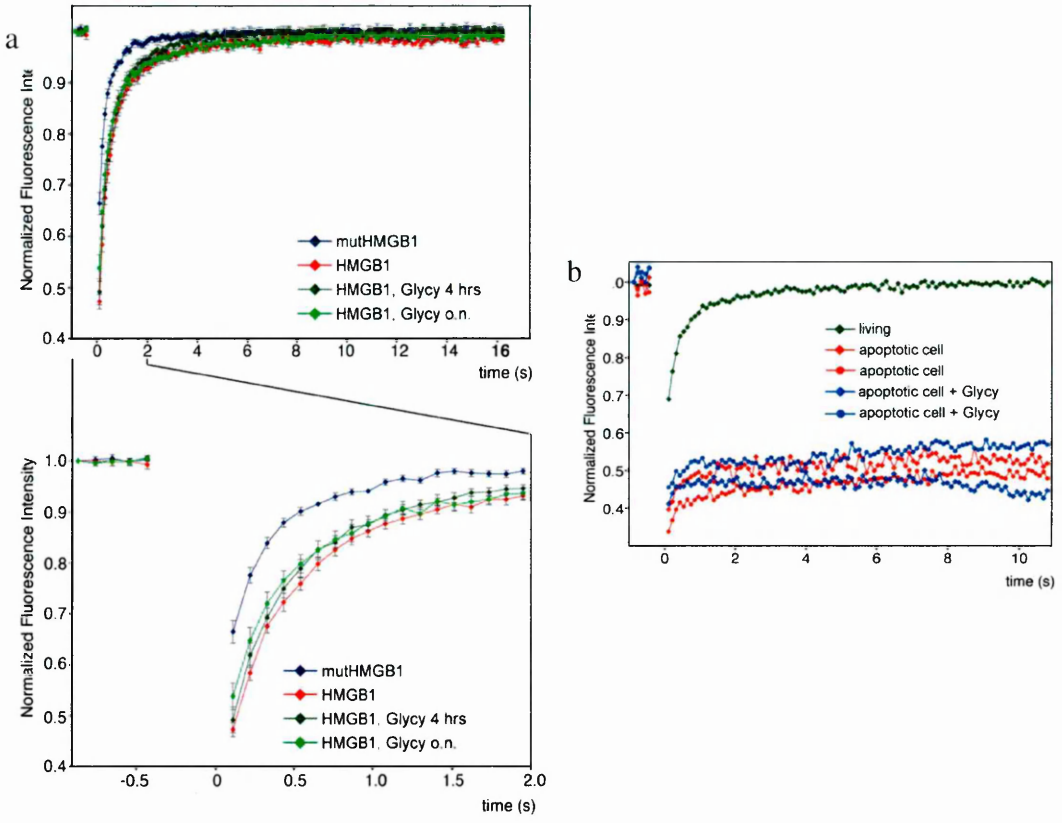


Fig. 1.14 a. FRAP analysis of HMGB1-GFP in living 3T3 fibroblasts incubated in the presence or absence of 100 mM glycyrrhizin for 4 hr or overnight. Compare for reference the mobility of an HMGB1 mutant unable to bind DNA (mutHMGB1). Dots indicate mean values, and error bars indicate standard error, from eight cells. Also shown is an enlargement of the recovery kinetics in the first 2 s. The effect of glycyrrhizin is statistically significant ($p < 0.01$ in paired two-way t test). b. FRAP analysis of HMGB1-GFP in 3T3 fibroblasts undergoing apoptosis for 16 hr after addition of 2 ng/ml TNF-a and 35 mM cycloheximide in the presence or absence of 100 mM glycyrrhizin. Only two cells per class are shown out of seven analyzed, to avoid clutter. The difference between treated and untreated apoptotic cells is not statistically significant (Mann-Whitney test). For reference, the recovery kinetics of HMGB1 in living cells is indicated.

of HMGB1-green fluorescent protein (GFP) increased a little after 4 hr ($p < 0.01$, ANOVA), and slightly more after overnight incubation, indicating that GL can indeed interfere with the binding of HMGB1 to DNA in living cells. However, the GL-induced increase in the mobility of intranuclear HMGB1-GFP was quantitatively modest, and far from the mobility of an HMGB1-GFP mutant that does not bind DNA (Agresti et al., 2005). This suggests that the effect of GL on the intranuclear functions of HMGB1 is modest. We also tested whether GL might interfere with the tight binding of HMGB1 to apoptotic chromatin (Scaffidi et al., 2002). Fibroblasts were induced to undergo apoptosis after exposure to TNF-a, in the presence or absence of 100 mM GL, and the mobility of HMGB1-GFP was measured by FRAP. HMGB1-GFP was immobilized both in cells treated and not treated with GL (Fig. 1.14).

1.7 Synopsis of the presented work

Starting from the observations about the possible interactions between triterpenic drugs and HMGB1 and moving from the results that indicate that a small molecule based therapy against HMGB1 is possible, we have investigated the nature of the binding of *GL* (and its derivative CBNX) to HMGB1 and provided an atomic level description of the protein-ligand complex. We adopted a hybrid experimental and computational approach that combined nuclear magnetic resonance spectroscopy, fluorescence spectroscopy and molecular dynamics and quantum mechanical calculations. The results of the thesis have been divided into two parts: a. the experimental results b. the structure calculations.

This work is intended as the first step towards the comprehension of the molecular mechanism that govern the effect of triterpenic drugs on HMGB1 induced inflammatory response and a necessary preliminary step for GL based drug design.

1.7.1 Synopsis of the experimental results

In the first part of the work (Chapter 3) the direct binding between protein and ligands is demonstrated, a K_d is reported (determined both by NMR and fluorescence spectrosopies) and the region of the protein that is involved in direct contact with the

ligand is reported as determined by the chemical shift difference (CSD) mapping method. Local mobility and overall diffusion for each HMGB1 Box in the free and bound forms has been studied by means of ^{15}N spin relaxation measurements, demonstrating the effect of the binding on the local stability and residual mobility of the protein on different timescales.

1.7.2 Synopsis of the structure calculations

In the second part of this thesis (Chapter 4) the problem of determining an atomic level description of the protein is faced: due to a quite high K_d it was not possible to treat the structure determination of the complex in a standard way (*i.e.*, distance restraints based structure determination), hence two different computational approach have been used for carbenoxolone (QRFF method) and for GL (data driven docking by HADDOCK) leading to a final model of the complexes that take into account all the experimental data collected on the system.

1.7.2.1 QRFF methodology

In the Quantum Refined Force Field methodology (Curioni et al., 2004) a set of molecular dynamics simulations is performed on the Box A - CBNX system starting from a structure obtained by standard (*i.e.*, using standard force fields) docking procedures based upon the knowledge of the ligand binding site and letting the system free to evolve towards minima. The system topology (*i.e.*, the ensemble of descriptors of molecular charges, bonds, plain angles and dihedral angles) has been built using a set of "protein tailored" parameters (*i.e.*, the GROMOS96 force field) for describing the inter- and intramolecular forces between the protein residues and between the protein, the ligand, the solvent and ions, adopting a quantum mechanical method to modify the parameters that describe the interactions of the ligand with the environment. This procedure is suggested by the different chemical nature of the protein and of the ligand that makes the representation of the latter less precise than the one of the former: due to the long-standing interest in protein simulation by means of molecular mechanics, available force fields for peptides have been frequently tested against experimental data

for a long time, hence they are more reliable in representing peptides dynamics than small molecules dynamics. This is also due to the relatively small chemical variability of proteins with respect to the huge variability of organic compounds, that in this sense are less prone to be parametrized.

The quantum mechanical method adopted is the plane waves scheme of DFT as implemented in the software CPMD (CPMD). The main feature of the QRFF procedure is the correction of the protein-based charges of the ligand: moreover, in the present work also a correction of dihedral angles is presented according to the calculation of vibrational spectrum of the CBNX molecule.

1.7.2.2 HADDOCK methodology

The molecule of main interest for the present study is GL. However, due to the difficult representation of sugars in the aforementioned molecular mechanics/dynamics scheme, the QRFF methodology cannot be easily implemented for the study of GL complexes with HMGB1 Box A and Box B. This led to the choice of a different method for the determination of the atomic level structure of these complexes: a docking scheme was adopted that could incorporate a set of experimental data that could "drive" the placing of the ligand on the surface of the macromolecule. This scheme is usually referred as data driven docking, and it is used as implemented in the HADDOCK package: the experimental data used as input are chemical shift displacement (CSD), ^{15}N spin relaxation, ^{1}H-^{15}N residual dipolar couplings (RDCs) and protein ligand NOEs. The whole set of experimental data are used in the early stages of docking (*i.e.*, rigid body docking) in order to find the correct relative "poses" of the two molecules that best satisfy the experimental information. The whole procedure is performed by a molecular mechanics engine based on the package CNS (Dominguez et al., 2003): after the initial stage that generates a large set of poses, the resulting structures are refined according to local sidechain flexibility and to the solvation of residues.

The HADDOCK based approach, with respect to the QRFF methodology, makes large use of experimental data to restraint the range of possibilities of protein-ligand poses: in this sense it is "logically" opposite to the previously presented approach, that

starts from an approximately determined position of the ligand and determines the best configuration of the complex only using first physical principles. The aim of the work is to determine the structure of the complex adopting different strategies according to the chemical difference and complexity of the molecule. Moreover, a comparison can be made between the two methods with respect to structural proteomics or drug discovery future projects.

2. Materials and Methods

2.1 Protein Expression and Purification

Recombinant HMGB1 constructs comprising Box A (residues 1–89), Box B (residues 90–175), and the full-length protein were cloned in a modified pET-24d vector (Novagen, Madison, WI) expressing a protein with an N-terminal 6-His tag, removable by cleavage with TEV protease. After expression and cleavage with TEV protease, the proteins have a residual N-terminal three-residue tag (GAM). Escherichia coli BL21 (DE3) cells were grown at 37°C until the optical density at 600 nm reached 0.8 absorbance units. Gene expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After 18 hr incubation at 25°C with shaking, cells were harvested by centrifugation. The cells were resuspended in lysis buffer (20 mM Tris-HCl (pH 8.0), 0.15 M NaCl, 10 mM imidazole, 2 mM β-mercaptoethanol, 0.2% NP40) and lysed by sonication. Cell debris was removed by centrifugation at 13,000 rpm for 30 min at 4°C in a microcentrifuge. The soluble 6-His-tagged HMGB1 protein was purified from the supernatant by affinity chromatography using Ni2+-NTA agarose resin (Qiagen, Hilden, Germany). After several (at least 4) washing steps, protein was eluted in 20 mM Tris (pH 8.0), 0.15 M NaCl, 0.3 M imidazole, 2 mM β-mercaptoethanol, and the 6-His tag was removed by overnight incubation at 25C with TEV protease. During incubation, the sample was dialyzed against 20 mM Tris (pH 8.0), 0.15 M NaCl, 2 mM bmercaptoethanol: the uncleaved 6-His-tagged protein and the TEV protease were then removed by repassing the sample over Ni2+-NTA resin. The protein sample was further purified on a HitrapQ ion-exchange column followed by gel filtration on a Superdex-75 column (Amersham Biosciences, Milan, Italy) equilibrated in 40 mM Tris (pH 8.0), 0.2MNaCl, 3mM dithiothreitol (DTT). All the purified HMGB1 polypeptides were exchanged into appropriate buffers for NMR (20 mM sodium phosphate (pH 7.3), 0.15 M NaCl, 5 mM DTT, 0.02% NaN3, 5% v/v D2O). Uniformly 15N- and 15N/13C-labeled proteins were prepared using M9 minimal bacterial growth media appropriately supplemented with

^{15}N-labeled ammonium chloride and ^{13}C-labeled glucose. Protein concentration was determined measuring the absorbance at 280 nm considering molar extinction coefficients of 9,186, 10,810, and 19,666 M^{-1} cm^{-1} for Box A, Box B, and HMGB1, respectively.

2.2 NMR Experiments

2.2.1 Ligands assignment

Carbenoxolone and Glycyrrhizin spectra were recorded at 298K on a Bruker Avance 500MHz spectrometer equipped with inverse triple-resonance probe and pulsed-field gradients, and processed with NMRPipe (Delaglio et al., 1995). Resonances were obtained and assigned from proton spectra with excitation sculpting scheme for water suppression (Hwang et al., 1995), 13C edited HSQC (Wilker et al. 1993), HSQC-DEPT (Palmer III et al., 1991), HMBC (Bax et al. 1986), COSY (Aue et al., 1976), TOCSY (Bax et al. 1985) and NOESY (Neuhaus et al., 2000). All the experiments have been performed on natural abundance samples at a concentration of 5 mM in 20 mM phosphate buffer (pH 7.3). The ligand numbering scheme is reported for Carbenoxolone in Fig.2.1; the same has been adopted for Glycyrrhizin triterpenic scaffold.

2.2.2 Proteins assignment and titrations.

NMR spectra of the free and bound protein were recorded at 298K on a Bruker Avance 600 MHz spectrometer (Karlsruhe, Germany) equipped with inverse triple-resonance cryoprobe and pulsed-field gradients. Spectra were processed with NMRPipe (Delaglio et al., 1995) and analyzed using XEASY (Bartels et al., 1995), PIPP (Garrett et al., 1991) and/or NMRView (Johnson et al., 1994). The 1H,13C,15N assignment of the backbone resonances of Box A, Box B, and full-length HMGB1 were obtained from HNCA (Witterkind et al., 1993), HNCACB and CT-HNCO (S. Grzesiek et al., 1992) with water-flip-back modifications (S. Grzesiek et al., 1993) and CBCA(CO)NH (S. Grzesiek et al., 1992) with WATERGATE scheme for water suppression (Piotto et al., 1992) using uniformly 15N,13C-labeled samples.

Typical protein concentration for NMR titration experiments was 0.3 mM in 20

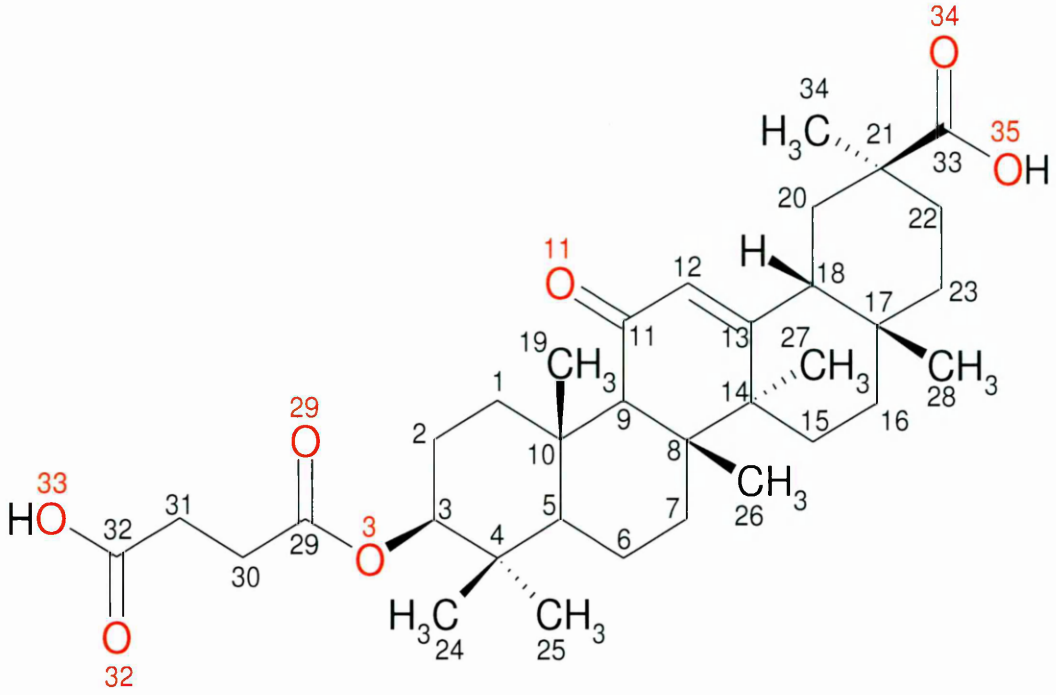


Fig.2.1 The chemical structure of Carbenoxolone. The enumeration adopted for Carbenoxolone is the same adopted for the triterpenic scaffold of Glycyrrhizin.

mM phosphate buffer (pH 7.3), 150 mM NaCl, 5 mM DTT. In order to minimize dilution and NMR signal loss, titrations were carried out by adding to the protein samples small aliquots of a concentrated (10 mM) ligand stock solution (glycyrrhizic acid monoammonium salt trihydrate, 98% purity (Acros Organics, Geels, Belgium), carbenoxolone disodium salt, 98% purity (Sigma, St. Louis, MO) in 20 mM phosphate buffer, 150 mM NaCl, at pH 7.3. For each titration point (typically 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4 equivalents of ligand), a two-dimensional water-flip-back ^{15}N-edited HSQC spectrum was acquired with 512 (100) complex points, 55 ms (60 ms) acquisition times, apodized by 60 shifted squared (sine) window functions, and zero filled to 1024 (512) points for ^{1}H (and ^{15}N), respectively. Assignment of the corresponding amide groups in the complexes was made by following individual crosspeaks through the titration series. For each residue, the weighted average of the ^{1}H and ^{15}N chemical shift difference was calculated (Fielding, 2003) as

$$CSD = \frac{1}{2}\sqrt{\frac{\Delta\delta^2 HN + \frac{\Delta\delta^2 N}{25}}{2}} \qquad (2.1)$$

Binding constants of glycyrrhizin to single Boxes (Box A and Box B) were estimated by monitoring the variation of CSD of individual peaks (eight peaks for each Box). Assuming a simple binary reaction between protein and ligand, dissociation constants were obtained from least squares fitting of CSD as a function of total ligand concentration according to the quadratic equation (Equation 1):

$$\delta_i = \frac{b - \sqrt{b^2 - 4ac}}{2} a \qquad (2.2)$$

with $a = (K_a/\delta_b)(P_t)$, $b = 1 + K_a((L_t) + (P_t))$ and $c = d_b K_a(L_t)$, where δ_i is the absolute change in chemical shift for each titration point, (L_{ti}) is the total ligand concentration at each titration point, (P_t) is the total protein concentration, $K_a = 1/K_d$ is the binding constant, and δ_b is the chemical shift of the resonance in the complex (Fielding et al., 2003). K_d and δ_b were used as fitting parameters using the Xmgrace program (http://plasma-gate.weizmann.ac.il/Grace/).

The proton chemical shifts of the bound glycyrrhizin and carbenoxolone were assigned by 15N/13C filter, 15N/13C-filtered TOCSY (60 ms) and 15N/13C filter, 15N/13C-filtered NOESY (120 ms) experiments. Intermolecular NOEs between glycyrrhizin/carbenoxolone and Box A were obtained from 2D NOESY experiments with $^{15}N/^{13}C$ filter in F2 (mixing time 70–120 ms) (Breeze, 2000); protein and ligand concentration were 0.6 mM and 0.9 mM, respectively.

2.2.3 Saturation Transfer Difference (STD) measurements

The STD measurements were performed on a Bruker DMX-500 MHz spectrometer employing a pulse scheme with WATERGATE water suppression and $T_{1\rho}$ filter (Mayer

and Meyer, 2001). The irradiation power (expressed in $\gamma B_1/2\pi$ units) was 20 Hz, which was applied on-resonance at 7 ppm or off-resonance at -12 ppm. The relaxation delay was set to 5 s to allow the ligands to relax between on and off-resonance irradiation cycles. A set of 7 spectra was collected for each complex, and the STD spectrum collected at each t_{sat} was the sum of 64 scans. The 10 spectra corresponded to the following saturation time: 0.2, 0.3, 0.6, 1.2, 1.8, 2.4 and 3.6 s.

2.2.4 Relaxation Measurements

All NMR experiments for the determination of ^{15}N spin relaxation were performed at 298K on a Bruker Avance 600 MHz (Karlsruhe) spectrometer with a cryogenic TCI probe. ^{15}N spin T_1, T_2, and $\{^1H\}$-^{15}N (heteronuclear) NOE (hetNOE) NMR experiments were performed for both the unliganded and liganded forms of both Box A and Box B (0.5 mM of ^{15}N labelled samples, pH 7.2, 20mM phosphate buffer) using sensitivity-enhanced pulse field gradient sequences (Farrow et al. 1994): heating effects were minimized in all spin relaxation experiments by collecting the data as interleaved pseudo-3D spectra. T_1 data were acquired with the following 14 relaxation delays: 20(duplicated for error analysis), 50, 100, 150, 200, 250, 370, 500, 650, 800, 1000, 1300, 1600, 2000 msec. T_2 data were obtained with the following 12 relaxation delays: 13.1 (duplicated for error analysis), 26.3, 39.5, 52.6, 65.7, 78.9, 105.1, 118.3, 131.1, 157.7, 184 msec. Both T_1 and T_2 experiments use a 2.5 seconds recycle delay. The $\{^1H\}$-^{15}N NOEs were measured by recording HSQC (Farrow et al., 1994) spectra with and without proton saturation: this spectrum has been recorded in an interleaved fashion using a 3 seconds recycle delay.

T_1, T_2 and hetNOE values have been obtained using the fitting routine implemented in the analysis program NMRView (Johnson et al.,1994) integrated with homemade scripts in Bash native language and in Tcl/Tk language for automation and efficient and fast data handling. The values have been extracted from the peak volume (V) decay curves using the two parameter exponential curve

$$V = V_{max} e^{-\frac{t}{T_2}} \qquad (2.3)$$

using a least squares fitting procedure with 100 iterations.

Correlation times τ_c that have been obtained from T_1 and T_2 using the empirical relationship:

$$\tau_c = \frac{1}{2\omega_N}\sqrt{\frac{6R_2}{R_1} - 7} \qquad (2.4)$$

Residues undergoing large amplitude fast motions have been detected by means of heteronuclear NOEs data (Tjandra et al., 1996). Residues (10-15% of the total amount) with too low values of NOE (< 0.65) have been excluded during the data analysis because of the weight they could have in the overall diffusional dynamics interpreted using a rigid body approach. Data have also been checked for residues that undergo slow motions due to chemical exchange related to anisotropic motion/difusion: they can be easily distinguished among the other low T_2 values through the error analysis of T_1 using the criteria that follow (Barbato et al., 1992):

$$\langle T_2 \rangle - T_{2,n} > \sigma \qquad (2.5)$$

$$\frac{T_{2,n} - \langle T_2 \rangle}{T_{2,n}} > 3(\langle T_1 \rangle - T_{1,n}) T_{1,n} \qquad (2.6)$$

The goodness of fit of T_1 and T_2 onto this diffusional tensor and of associated errors has been analyzed using a 500 steps Montecarlo simulated annealing procedure as implemented in the software TENSOR2 starting from the predicted best fit of diffusional tensor. An F test allowed the discrimination between a fully symmetric, axially symmetric or fully asymmetric model for the description of the overal diffusional motion of the protein (Dosset et al., 2000).

Lipari-Szabo "model-free" analysis of relaxation data has been performed with the help of the program TENSOR2 (Dosset et al., 2000).

2.2.5 ^{15}N-^{1}H Residual Dipolar Couplings (RDCs)

The ^{15}N-^{1}H RDC measurement were performed adopting the method of acrylamide cylindrical gels axial compression described by Grzesiek and coworkers (Sass et al., 2001): this alignment medium was chosen because of the highly negative charge of HMGB1 that disallowed the usage of phages (a negatively charged alignment media) due to strong electrostatic interactions that lead to pronounced signal broadening and/or peak disappearance (only a total amount of ~ 10 signals resulted detectable considering the the two Boxes). A cylindrical plastic tube of suitable 3.5 mm was closed on one end by parafilm. An acrylamide solution was prepared at the final percentage of 6% from a stock of 29.2% w/v acrylamide and 0.78% w/v N,N0 -methylenebisacrylamide. 250uL of the 6% acrylamide solution was filled into the tube, just after the polymerization started by the addition of 0.1%w/v ammoniumpersulfate and 0.5% w/v tetramethylethylenediamine (TEMED). After polymerization was complete, the gel was pushed out from the tube by the application of gentle air pressure by means of a syringe: at this stage the gel presented a legth of 22 mm. After washing 5 times the gel in 40mL of MilliQ water for 10 minutes under gentle shaking it was dried over night on a smooth plastic support at 37°C in a drying oven. The dried gel was then placed into an NMR shigemi sample tube and 310μL of buffer solution containing up to 0.3 mM of protein or complexes (Box A or Box B alone or complexed with GL) was then added. After 3-4 h of swelling, the gel had filled out the sample tube and high resolution NMR measurements could be performed after axial pressure exerted by the plunger until the gel reached a length of 15 mm. The whole procedure is summerized in gel and tube pictures in Fig.2.2. The RDCs measurements were performed on a Bruker DMX-500 MHz spectrometer: a sensitivity-enhanced HSQC experiment that separates the upfield and downfield components in the indirect dimension into different subspectra with a sequence similar to the generalized TROSY scheme and decoupling of the N-nucleus during detection was performed for each sample (Cordier et al., 1999). The same kind of

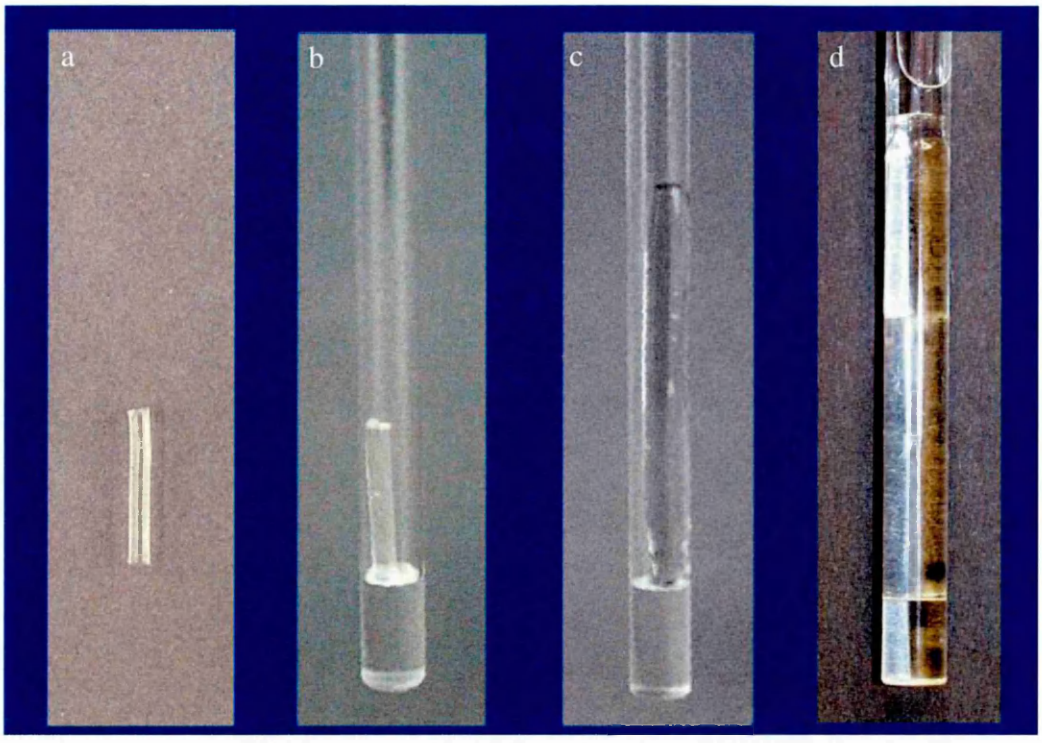


Fig.2.2 Acrylamide gel cylinder preparation for the RDCs measurements in a shigemi tube: a. the polymerized acrylamide gel cylinder; b. the gel cylinder dried and inserted in the shigemi tube; c. the gel during swelling with the protein solution; d. the gel completely rehydrated and pressed by the shigemi plunger. Real dimensions are reported on the picture.experiment was repeated for simple solutions of proteins and complexes without acrylamide gels in order to get a set of reference spectra.

All the data have been first visually inspected with the help of MODULE (Dosset et al., 2001) and analyzed with the DC program implemented in nmrPipe suite (Delaglio et al, 1995). The Q factors (Cornilescu et al, 1998) with respect to all the the available structure were computed according to the following formula:

$$Q = \sum_{1}^{N} \sqrt{\frac{DC_{calc} - DC_{obs}}{N D_a^2 (4 + 3R^2)/5}} \qquad (2.7)$$

where DC_{calc} and DC_{obs} are respectively the dipolar couplings calculated (i.e., from a

structure provided) and observed for each residues, N is the number of DC values considered ,$D_a = D_{zz}/2$ is the axial component of the order tensor $\boldsymbol{D}$ expressed in terms of its eigenvalues and $R = 2(D_{xx} - D_{yy})/3D_{zz}$ is the rhomboedricity of the tensor.

2.3 Docking, Molecular Mechanics and Quantum Mechanics Calculations

2.3.1 Protein only simulations

All the molecular dynamics simulations on HMGB1 Boxes A and B have been performed using GROMACS 3.2 and GROMACS 3.3 (Berendsen et al., 1995) the force field adopted was the implemented GROMOS 43a1 (van Gunsteren et al., 1996). The calculations were performed on a 56 CPUs IBM eServer BladeCenter equipped with Myrinet communication system.

Box A and Box B structures used for performing MD simulations were obtained as average structures of the bundles of structures deposited on the Protein Data Bank with accession ID 1aab and 1hmf respectively (Weir ert al., 1993; Hardman et al., 1995): the structures included in the bundles have a backbone RMSD (computed with respect to the average structures) of 0.8 ± 0.2 Å and of 0.9 ± 0.2 Å for Box A and Box B respectively. The protonation state considered for the simulations was for both the HMGB1 Boxes the one corresponding to pH 7, i.e., with protonated histidines and with Asp, Glu, Lys and Arg charges fixed at their native values, consistent with the neutral environment. Starting systems for the MD simulations were built with the help of the Visual Molecular Dynamics visualizer (VMD) (Humphrey et al., 1996) in order to a. place the protein in the geometric center of the simulation box; b. ensure enough distance between the surface of the protein and the Box edges, i.e. in order to ensure no interactions of the protein with its periodic images in the combined PBC/PME treatment (see further). The systems built in this way consisted of the protein, water molecules (~ 17100 in both cases) and counterions (7 chloride ions for Box A and 6 for Box B) added in order to balance the system net charge. Proteins and water were coupled separately to a temperature bath at 300 K, using a Berendsen thermostat (Berendsen et al., 1984) with

a coupling constant t 0.1 ps. In all simulations the pressure was kept at 1.0 bar using isotropic pressure coupling (Berendsen et al., 1984) with p 1 ps, hence obtaining an isocoric-isobaric simulation ensemble. Bond lengths were constrained using the LINCS algorithm (Hess et al., 1997) and Lennard–Jones interactions were calculated with a 0.9/1.4 nm twin-range cutoff. The short-range electrostatic interactions were calculated to 0.9 nm, and the Particle Mesh Ewald (PME) algorithm was used for the long-range interactions. All trajectories were calculated in periodic rectangular Boxes of explicit simple point charge (SPC) water molecules (Berendsen et al., 1981). Periodic boundary conditions were applied in all the three cartesian dimensions. Solvent was equilibrated by 500 ps of solute position restrained MD with a force constant of 1000 kJ mol^{-1} nm^{-1}, followed by 10 ns of unrestrained MD with a 2 fs time step.

In order to obtain an improved sampling of the phase space of the proteins (especially in more flexible regions) and to overcome any local minimum present on potential energy surface of the protein in solution, a Replica Exchange Molecular Dynamics scheme (Mitsutake et al. 2001) has also been adopted for both the systems in NPT ensemble. For Box A and Box B REMD simulations with 16 replicas (with a simulation time of 5ns for each replica) have been performed: temperature difference between neighboring replicas was set between 2 and 4 K, ensuring an exchange rate between 10 and 20% for each pair of neighboring replicas. The simulation parameters adopted, for sake of comparison, were the same indicated above for "standard" molecular dynamics simulations of Box A and B.

2.3.2 Box A-Carbenoxolone complex simulations

Molecular dynamics simulations have been performed in order to describe the dynamical behaviour of the Box A – Carbenoxolone complex in solution and find a relationship with the experimental data. Due to the incorrect representation of small ligands using protein-tailored force fields in the most common MD codes, especially for what concerns electrostatics, a more detailed treatment of ligand charges is required.

We created the ligand (i.e., carbenoxolone) GROMOS topology using the PRODRG (Schuettenkolpf et al., 2004) web based software considering deprotonated

oxygen atoms (i.e., charge state of the molecule was +2). These were later modified according to the QRFF method (Curioni et al., 2004), which produces quantum-refined charges consistent with the original parameterization, thus maintaining the balance between electrostatic and van der Waals forces. The central step of this procedure consisted of *ab initio* calculations of the ligand and of its electron density based on density-functional theory (DFT) as implemented in the Car-Parinello Molecular Dynamics package (CPMD). This specific DFT implementation used the Becke–Lee–Yang–Parr (BLYP; Becke, 1988) exchange-correlation functional, norm-conserving pseudopotentials (Troullier et al., 1993) and a plane-wave basis set with a 70 Ry cutoff.

First the structure of CBNX (Fig.2.1) was extracted from its complex with 3 alpha, 20 beta-hydroxysteroid dehydrogenase (Ghosh et al., 1994) and optimized within DFT–BLYP, starting from one reported for this ligand in another complex . The changes of bond-lengths with respect to these experimental data were of the order of 0.01 Å with the exception of the C–O double bond which was 0.1 Å larger. The refined values of the effective charges (given in Fig. 2.3a) were then derived from a restrained fitting (Curioni et al., 2004) to the DFT electrostatic potential as in previous applications (Ferrara et al., 2006). In particular, for the atoms not included in the GROMOS parameterization (i.e., hydrogen atoms), the restraint value of the effective charges was set to zero.

To ensure the transferability of the QRFF, the final values were obtained as averages over 20 distinct conformations of the CBNX extracted from a 1 ns trajectory (every 50 ps) of GROMOS-based MD simulations of the ligand–protein complex in water: the simulation was performed using the software package GROMACS 3.3 and the parameters used for this simulation were the same indicated in the previous subsection (c1). CBNX configurations did not change substantially over the whole trajectory, thus revealing a certain rigidity of the ligand in the complex and resulting in a maximum deviation of 5% on the effective charges. However, as expected, they were very different from that of the molecule at equilibrium in vacuum. The average values, which were used in the calculations, are reported in Table 2.1 and compared to the unrefined ones.

The deviation from the unrefined values is shown in Fig. 2.3b for all atoms.

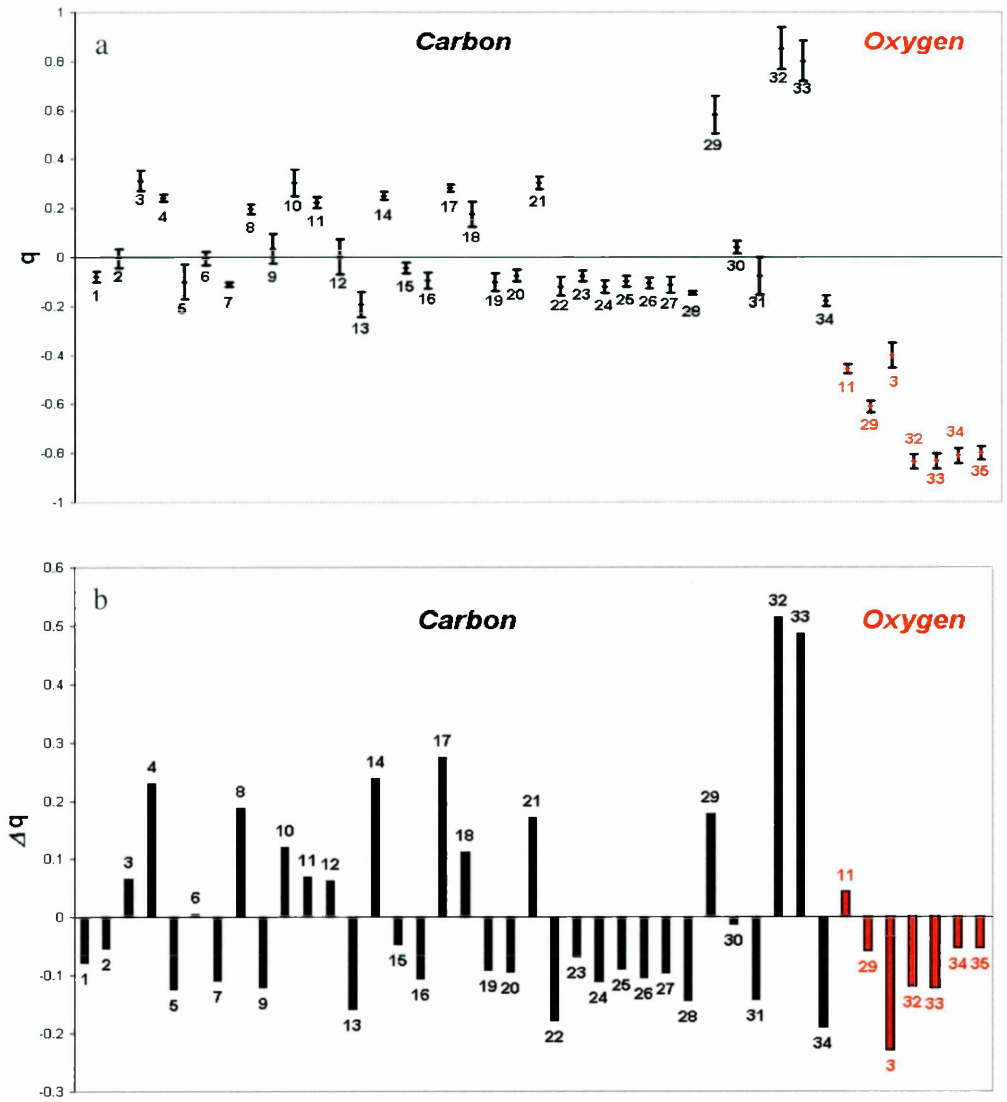


Fig. 2.3 a. Charges distribution of Carbenoxolone after QRFF refinement of charges; b. differences between the new values obtained by QRFF and the values provided by the PRODRG server.

GROMOS Atom Type	Atom Name	charges PRODRG	charges QRFF
CH2	C1	0.000	-0.078
CH2	C2	0.049	-0.005
CS1	C3	0.242	0.308
CH1	C4	0.000	0.231
CH1	C5	0.000	-0.125
CH2	C6	0.000	0.006
CH2	C7	0.000	-0.109
CH1	C8	0.000	0.188
CH1	C9	0.146	0.025
CH1	C10	0.198	0.318
CB	C11	0.157	0.227
R61	C12	-0.036	0.027
CB	C13	-0.052	-0.210
CH1	C14	0.000	0.238
CH2	C15	0.000	-0.047
CH2	C16	0.000	-0.107
CH1	C17	0.000	0.274
CH1	C18	0.088	0.200
CH3	C19	0.001	-0.090
CH2	C20	0.000	-0.095
CH1	C21	0.139	0.310
CH2	C22	0.039	-0.140
CH2	C23	0.000	-0.069
CH3	C24	0.000	-0.111
CH3	C25	0.000	-0.089
CH3	C26	0.000	-0.104
CH3	C27	0.000	-0.097
CH3	C28	0.000	-0.144

C	C29	0.434	0.612
CH2	C30	0.041	0.028
CH2	C31	0.041	-0.102
C	C32	0.364	0.879
C	C33	0.346	0.833
CH3	C34	0.000	-0.190
OS	O3	-0.160	-0.390
O	O11	-0.502	-0.458
O	O29	-0.565	-0.624
OM	O32	-0.723	-0.843
OM	O33	-0.723	-0.846
OM	O34	-0.762	-0.816
OM	O35	-0.762	-0.816

Table 2.1 Values of the effective atomic charges in the PRODRG and in the QRFF parametrization.

Clearly the refinement is active all over the molecule, but the major changes pertain to the carbon atoms of the carbonyl endgroups. The major change relative to the PRODRG-derived values was found for the C(11)–C(12) bond consistent with its partial double-bond character. Moreover the value of an improper dihedral term, involving C(11), C(12), C(13) and O(11), that was missing in the PRODRG parameterization, was introduced (as an additional force constat) as obtained from our DFT calculations (2448.6 kJ nm^2 mol^{-1}).

The initial step for the identification of the binding site was made on the basis of the information derived from chemical shift perturbation data and of a standard docking procedure using AUTODOCK (Morris et al., 1998). The system for the molecular dynamics run was setup as previously indicated and the same parameters have been

used: the simulation box was first minimized using a steepest descent algorithm and the solvent was equilibrated separately, afterwards it was allowed to evolve in an NVT ensemble until it stabilized itself within the last 2 ns of the run. To verify of our results we employed the replica-exchange MD method (REMD) (Mitsutake et al. 2001) and ran simulations keeping the number of particles fixed in the isocore–isothermal ensemble (66 ns). The replicas covered the range between 300 and 370 K, with an interval of 2–4 K. A further attempt to verify our predictions with an analogous protocol was done adopting an isobaric–isothermal ensemble (11 ns).

For the sake of comparison, a simulation has been performed with the same protocol as above using the unrefined FF. Again, a subsequent REMD simulation (in the interval 300–370 K for 60 ns total time) was performed in order to verify the validity of our results.

2.3.3 Experimentally driven docking calculations of HMGB1-GL complexes

Molecular docking of glycyrrhizin on HMG Boxes was performed using the software HADDOCK 2.1_devel, which makes use of chemical-shift perturbation data to drive the docking while allowing various degrees of flexibility (Dominguez et al. 1998). The protocol follows a three-stage docking procedure which includes (1) randomization of orientations and rigid body minimization, (2) simulated annealing in torsion angle space, and (3) refinement in Cartesian space with explicit water. Residues showing chemical shifts larger than the average chemical shift difference plus one standard deviation were used to define active residues according to the HADDOCK definition, and were used as ambiguous interaction restraints (AIRs) (Dominguez et al., 2003). Residues that were close to the threshold, or with small (<50%) solvent-accessible surface areas, were not included in AIRs. Passive residues coincided with the region which was allowed to be flexible during the semiflexible simulated annealing step (Dominguez et al., 2003).

Intermolecular nOes, R_1/R_2 values for each residues and ^{15}N-^{1}H RDCs values for each backbone amide bond were included as direct restraints in the calculations only in the final run: intermolecular NOEs were treated as unambiguous distance restraints, R_1/R_2 values have been obtained from relaxation measurements using the diffusion

anisotropy model of TENSOR2 (Dosset et al., 2000; Van Dijk et al., 2006) and RDCs have been transformed into intervector projection angle restraints (Meiler et al. 2000, Van Dijk et al., 2005). Optimized parameters for liquid simulation (OPLS) were used for the protein. The geometric coordinates and parameters for the ligand were calculated and optimized using the PRODRG server (Schuettenkolpf et al., 2004). During the rigid body docking, 2000 structures were calculated. The best 200 solutions in terms of intermolecular energies were selected for a semiflexible simulated annealing, followed by water refinement. The analysis of the simulations was performed applying in-house python and tcl scripts. Root-mean-square deviations (RMSD) were calculated using the ProFit program (http://www.bioinf.org.uk/software/profit/). The fitting of the protein was performed on the residues belonging to the binding site using the McLachlan algorithm (McLachlan, 1982); the RMSD of glycyrrhizin was calculated only on the heavy atoms of the triterpene scaffold. The final RMSD matrix was then clustered using the algorithm described in Daura et al. (Daura et al. 1999), where a cluster is defined as an ensemble of at least two conformations displaying an RMSD smaller than 0.8 Å. The final structures were clustered and scored using a combination of energy terms (total energy, intermolecular van der Waals and electrostatic energies, restraint energies).

2.4 Fluorescence Measurements

Experiments were performed on a Varian Eclipse instrument (Victoria, Australia) with excitation at 273 nm (slit width 5 nm) and emission at 325 nm (slit width 5 nm). The initial protein concentration was 3 μM; after each ligand addition, the samples were left to equilibrate for 5 min. The ligand concentration at the end of the titration was 250 mM for Box B and 270 mM for Box A and HMGB1; the added ligand solution did not exceed 10% of the initial volume. At each data point, the protein concentration was corrected for the dilution due to ligand addition. All measurements were performed at 25°C in 20 mM phosphate buffer (pH 7.2), 150 mM NaCl. Dissociation constants (K_d) for Box A and Box B were estimated considering a single-step binding process with an equimolar stoichiometry.

The free ligand concentration in solution was considered approximately equal to the total concentration of added ligand. Experimental data were fitted by nonlinear regression (Xmgrace) to the equation:

$$F = F_0 - \frac{F_0 - F_\infty}{[L_0] + K_d}[L_0] \qquad (2.8)$$

where F is the fluorescence intensity measured at each titration point, $[L_0]$ is the total ligand concentration added at each point, F_0 is the initial fluorescence intensity of the unbound protein, and F_∞ corresponds to the fluorescence intensity of the bound final species.

K_d values for HMGB1 were estimated considering a mechanism in which two ligand molecules bind simultaneously and independently to Box A and Box B. Experimental binding curves have been fitted through the following equation:

$$F = F_{\cdot,A} - \frac{F_{\cdot,A} - F_{\infty,A}}{[L.] + K_{dA}}[L.] + F_{\cdot,B} - \frac{F_{\cdot,B} - F_{\infty,B}}{[L.] + K_{dB}}[L.] \qquad (2.9)$$

Notably, the calculated K_{dA} and K_{dB} were very similar; for this reason, we then used a model where $K_{dA} = K_{dB} = K_{d,app}$ and therefore the folowing equation has been used:

$$F = F_{0,FL} - 2\frac{F_0 - F_\infty}{[L_0] + K_{d,app}}[L_0] \qquad (2.10)$$

3. The interaction of HMGB1 with Glycyrrhizin and Carbenoxolone

3.1 Summary

In order to characterize the binding of Glycyrrhizin (GL) to HMGB1 Box A and Box B, NMR and intrinsic triptophane fluorescence based titrations with GL were performed. NMR experiments also allowed the identification of the residues that are mostly perturbed by the presence of the ligand. Moreover, they both allowed the determination of the K_d of the process. Relaxation measurements have been performed in order to get a detailed description of the local and overall mobility of the protein upon binding. These data, taken all together, are the first step towards the atomic-level comprehension of the nature of the GL-HMGB1 binding. This part of the work contributes to the understanding of the molecular mechanism that regulates part of the cytokine activity of HMGB1 and identifies in GL a possible inhibitor of HMGB1 mediated cell migration and proliferation.

3.2 Results

3.2.1 Ligands and protein resonance assignment

NMR assignment of GL and carbenoxolone (CBNX) have been obtained starting from chemical shift datasets reported in literature and using the 1D and 2D NMR experiments described in the previous chapter. As expected, the spectra of the two compounds appeared to be very similar due to the presence of a common triterpenic scaffold, and the glycosidic moiety of GL has been easily recognized in the spectra and assigned (see Fig. 3.1). The two compounds displayed a different solubility: GL was soluble up to 50 mM in phosphate buffer 20 mM at pH 7.4, whereas CBNX had a lower solubility (less than 10 mM) and was very unstable in solution (precipitation occurred within some days). The unassigned peaks were assigned to impurities generated during the process of chemical extraction of the compound from plants (Tianwei et al. 2002).

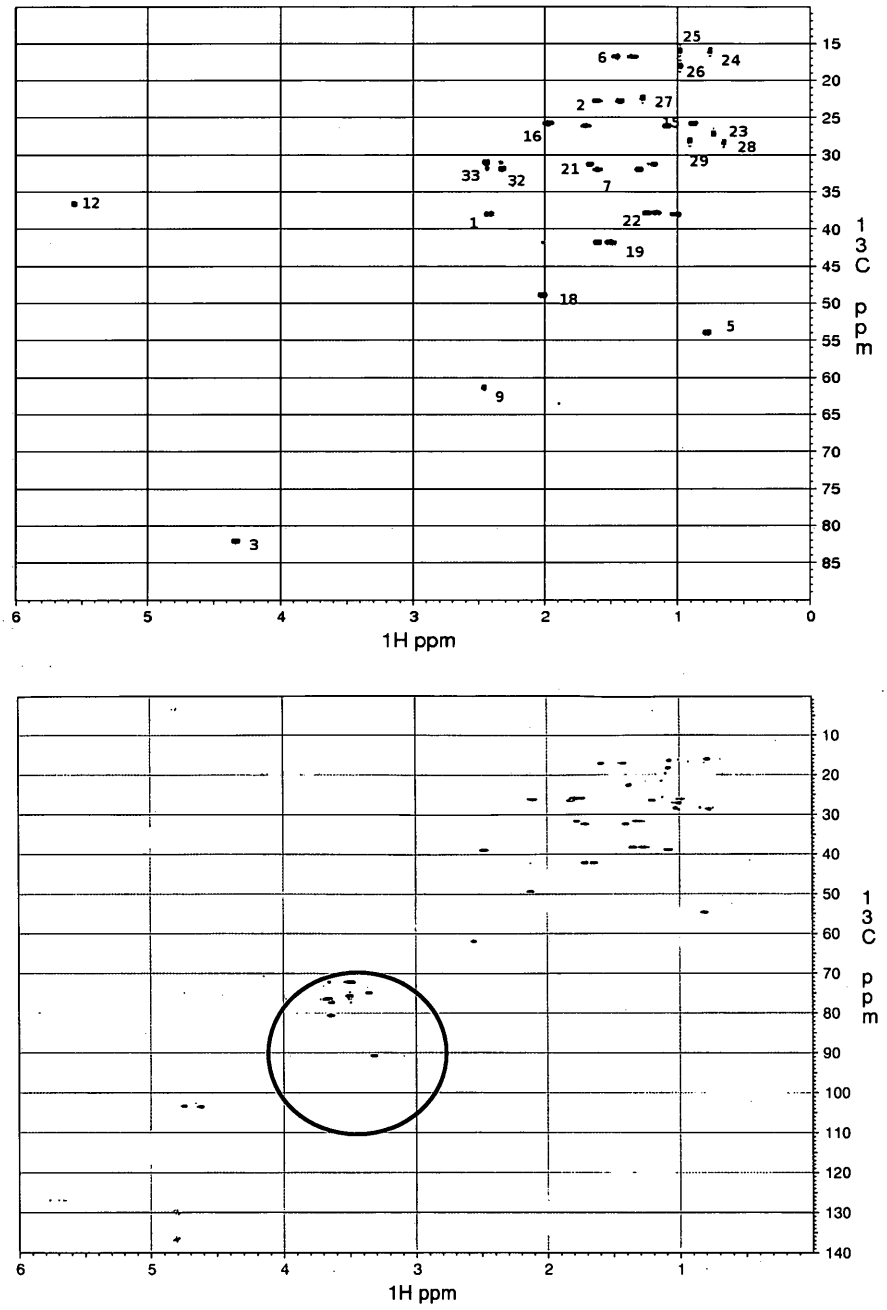


Fig. 3.1 CBNX (a) and GL (b) ^{1}H-^{13}C HSQC spectra recorded in isotopic natural abundance. The signals due to ^{1}H-^{13}C coherences typical of sugar rings are circled in red on the spectra of GL.

The sequential backbone assignment of full-length (FL) HMGB1 and its fragments Box A and Box B have been obtained using standard triple resonance

experiments (Sattler et al, 1999) performed on the triple resonance spectra described in the previous chapter. The resonances have been deposited in the BMRB data bank (http://www.bmrb.wisc.edu) with accession numbers 15148 and 15149.

3.2.2 GL binds directly to HMGB1 Boxes.

Direct binding of GL to HMGB1 was probed using NMR chemical-shift differences (CSD), a highly sensitive tool for proving interactions and for mapping binding sites and detecting residues which directly interact with the ligand or that are indirectly affected by the association. Two-dimensional ^{1}H-^{15}N HSQC spectra of full-length ^{15}N-labeled HMGB1 were recorded to monitor the changes in the ^{1}H-^{15}N chemical shifts of the backbone amide groups induced by successive additions of GL. A comparison of the spectra in the absence and presence of a 4-fold excess of GL is shown in Fig. 3.2a. The complex is in fast exchange on the chemical-shift timescale, as demonstrated by the monotonic trend of chemical shift displacements with respect to the amount of ligand added for each point of titration (an example is reported in Fig. 3.2b); this facilitated the assignments of the resonances in the complex, which were obtained following the cross peaks of HMGB1 upon addition of increasing amounts of GL.

3.2.3 The ligand binding surface is identified by means of chemical shift mapping

To identify the interaction surface, the average chemical-shift changes between the free and the bound state were plotted versus the HMGB1 residue numbers (Fig. 3.3). Most of the protein residues do not experience relevant chemical-shift perturbation, indicating that GL does not alter the overall protein structure. The global preservation of the secondary structure was further confirmed by circular dichroism (CD) spectroscopy (Fig. 3.4). Resonances with significant CSD (deviating more than one standard deviation from the mean CSD) included residues Phe17, Gln20 (side-chain protons), Arg23, Gln25, Lys43, and Cys44 for Box A (Fig. 3.3a and 3.5) and Arg109, Ile112, Asp123, and Ala125 for Box B (Fig. 3.3a and 3.6). No interactions were observed with the linker

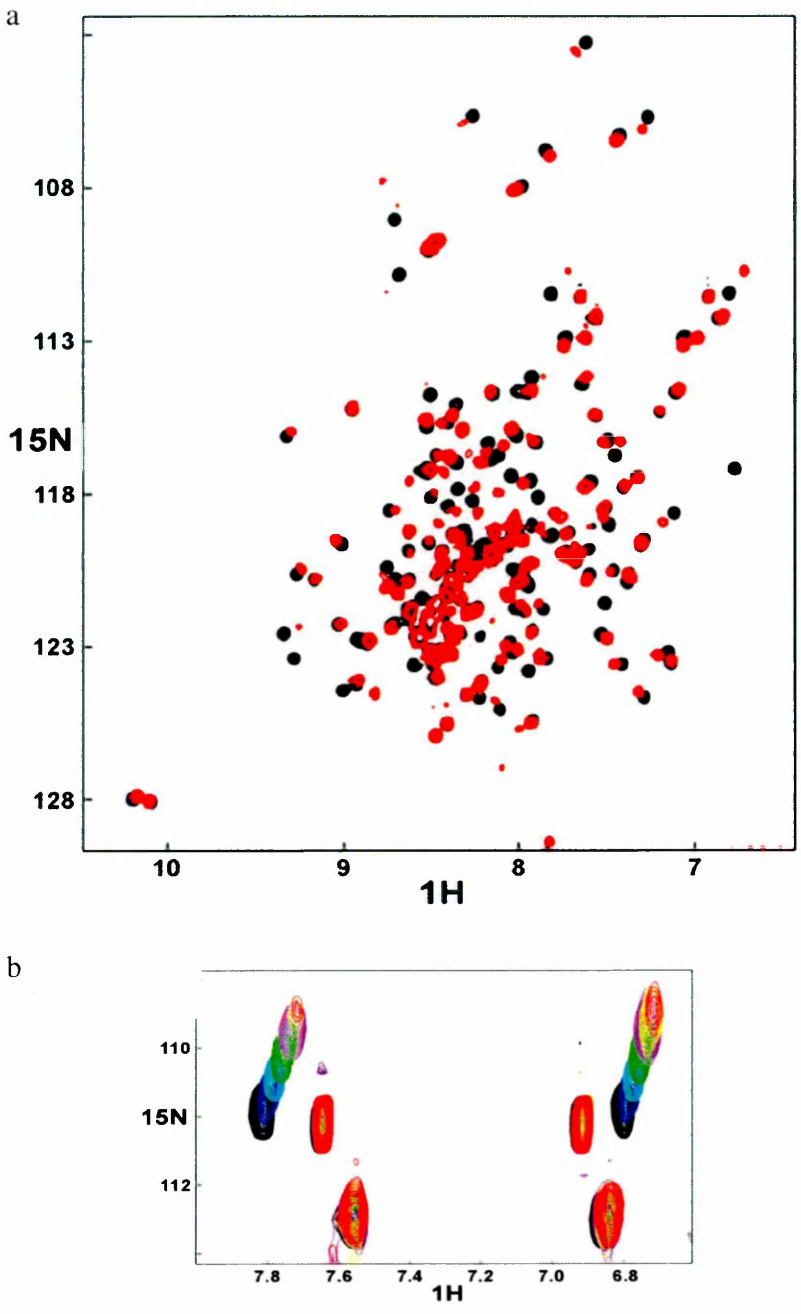


Fig. 3.2 (a) A comparison of the spectra of 15N labeled FL HMGB1 in the absence (black) and presence (red) of a 4-fold excess of glycyrrhizin. (b) A selected region (Gln20 sidechain) of HMGB1 spectra during the titration with glycyrrhizin (0.5, 1, 1.5, 2, 3, 4 equivalents of ligand). The starting and end points of the titration are represented in black and red, respectively. The largest shifts were observed for the amino group of Gln20.

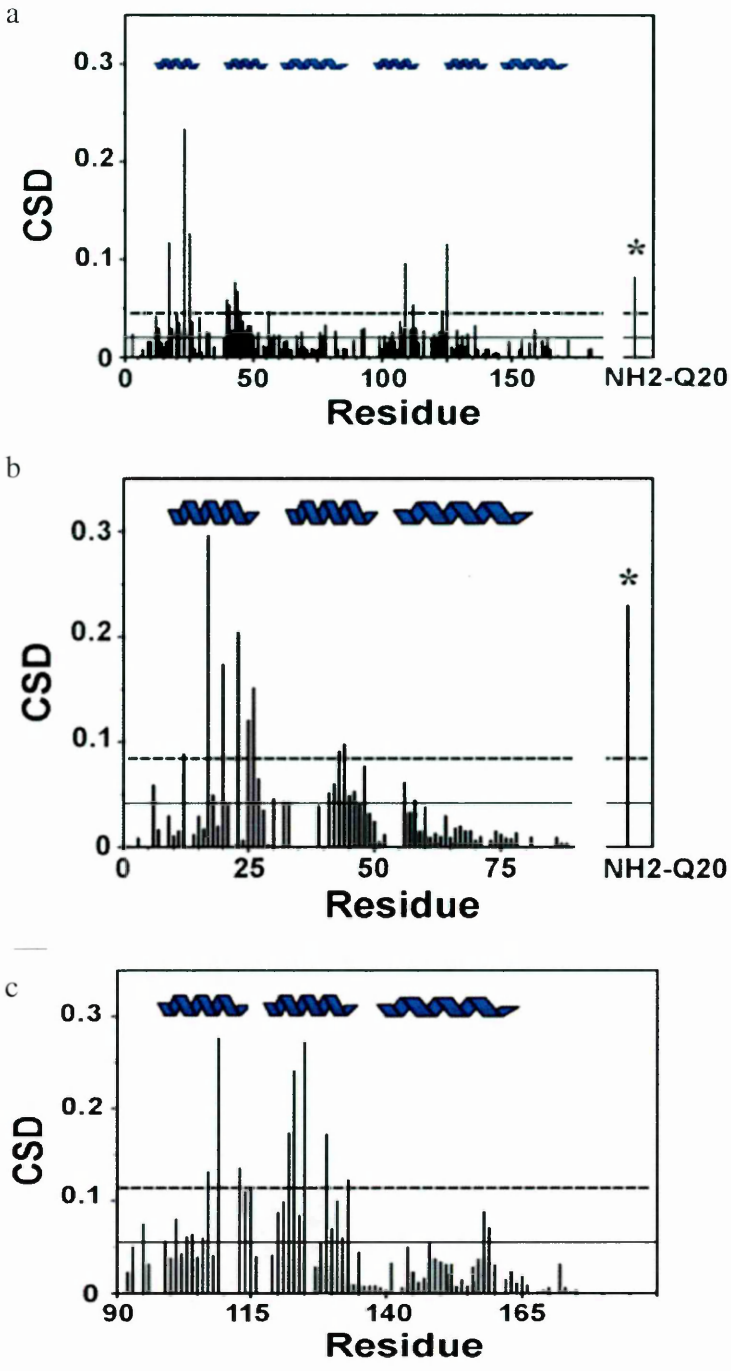


Fig. 3.3 Histogram representation of chemical shift differences (CSD) between GL bound (protein – ligand ratio 1:4) and free state of FL HMGB1 (a), Box A (b) and Box B (c). Secondary structure elements are reported on the top of the images. Gln20 sidechain are marked by an asterisk and reported with an increased order number in (A) and in (B). CSDs have been computed as reported in Materials and Methods section.

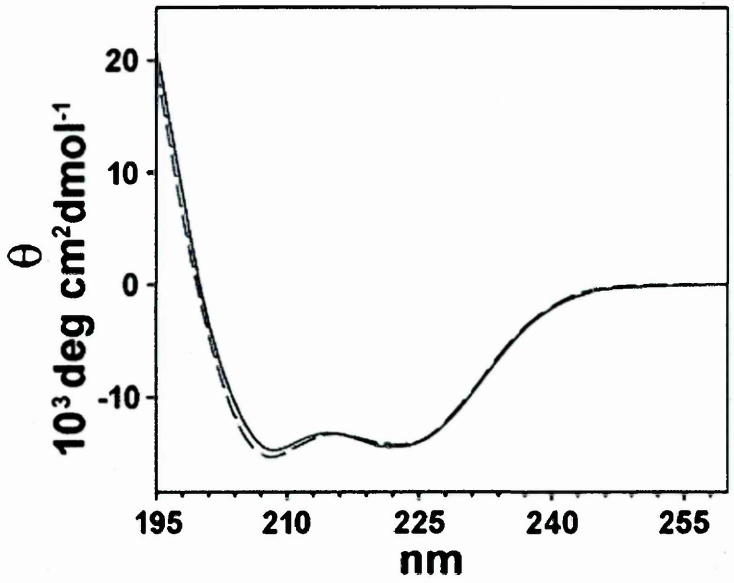


Fig. 3.4 Circular Dichroism spectrum of free (continuous line) and GL bound (dashed line) HMGB1 FL protein.

region between the two boxes and the C terminal region. The individual peaks corresponding to the acidic tail residues could not be assigned because of high spectral overlap and exchange with the solvent (Knapp et al., 2004); however, these peaks did not shift upon addition of GL, excluding their involvement in the binding.

Previous structural studies have demonstrated that the two HMG boxes behave as rigid independent domains that do not interact with each other in the context of the full-length protein (Knapp et al., 2004); it was therefore verified whether GL binds similarly to the two individual HMG boxes (^{15}N-labeled box A and box B) (Fig. 3.3b and 3.3c). Indeed, residues mostly affected by the presence of the ligand all clustered on the first two helices of both HMG boxes, and coincided with the residues showing high CSD in the full-length protein, indicating that the two Boxes actually behave in a non-cooperative way. Additional residues located on Helices 1 and 2 (His26, Lys42, Gln107, Lys113, Gly122, Gly129, Met133) were more affected in the isolated domains than within the entire protein. This effect might be due to a slightly different orientation of the ligand, or to small structural differences of the entire protein when compared to the

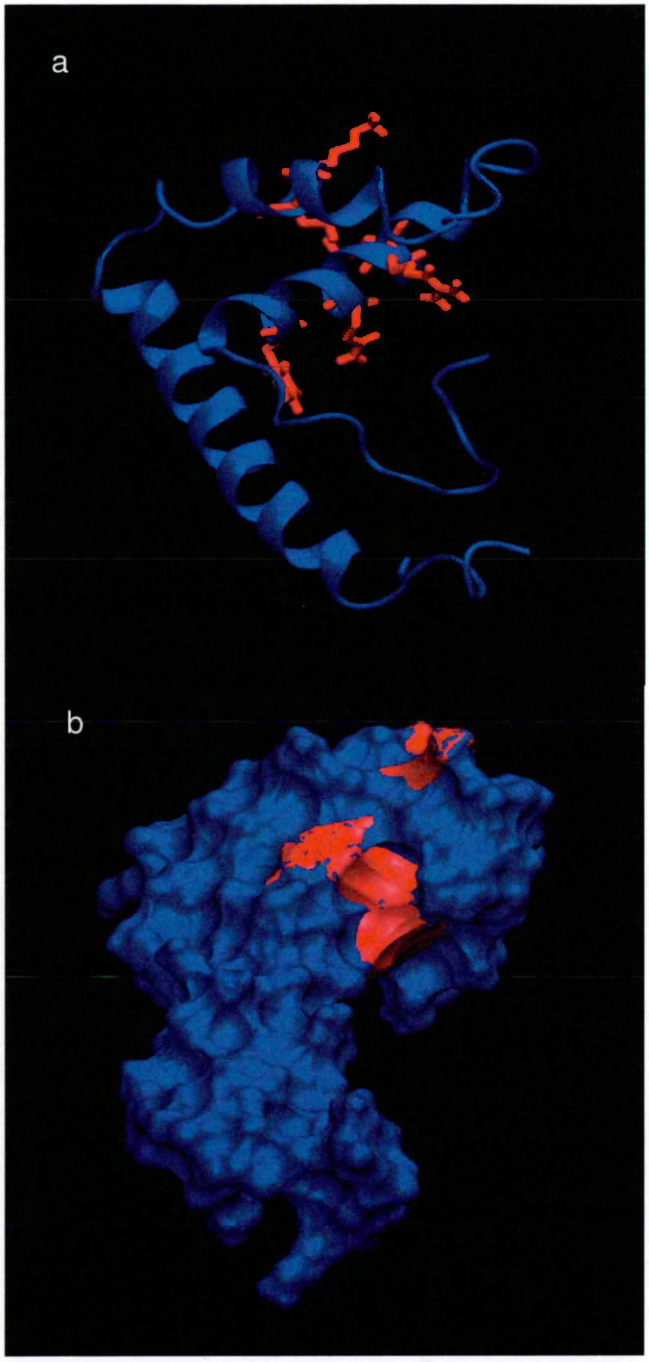


Fig. 3.5 GL binding site of Box A. Residues that have been characterized by the biggest displacement in CSD (17, 20, 23, 25, 43 and 44) are highlighted in red. Cartoon (a) and surface (b) representations are reported. The protein in (b) has been slightly reoriented in order to show almost completely the binding surface.

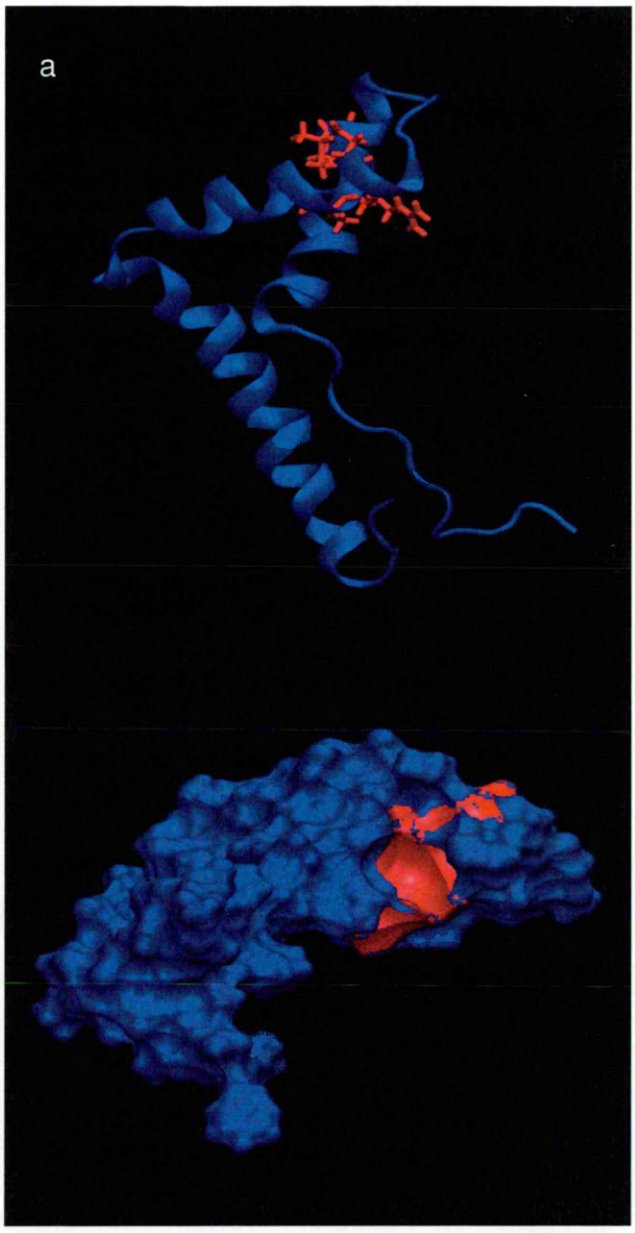


Fig. 3.6 GL binding site of Box B. Residues that have been characterized by the largest displacement in CSD (109, 112, 123, 125) are highlighted in red. Cartoon (a) and surface (b) representations are reported. The protein in (b) has been slightly reoriented in order to show almost completely the binding surface.

isolated domains. On both HMG boxes (whether in HMGB1 or in isolation), the binding site is located at the crux of the typical L-shape fold, which has a small solvent-exposed hydrophobic surface suitable for favorable van der Waals interactions with the triterpene

scaffold of GL (Fig. 3.5 and 3.6). Interestingly, in the upfield region of the monodimensional proton spectrum of both HMGB1 and Box A, variations in chemical shifts are also observed for the well-resolved ^{1}H methyl resonances of Val19 and Val35, located on Helices 1 and 2 of box A, further suggesting an involvement of this region in the binding (Fig. 3.7). The titration experiments also indicate that Helix 3 is not involved in the binding. We repeated the titrations with carbenoxolone: in the case of Box A, we found that the highest chemical-shift perturbations of the backbone ^{1}H and ^{15}N weretrans again located on the first two helices, as observed for GL (Fig. 3.8). Unfortunately, the spectra for Box B and the full-length protein at higher carbenoxolone concentrations were of low quality, and showed evidence of aggregation. This suggests that the triterpenic moiety of the ligands drives the binding, whereas the glycosidic moiety of GL contributes to its greater solubility with respect to carbenoxolone: this can explain also the partially different solubility of the complexes in molar excess of the ligand.

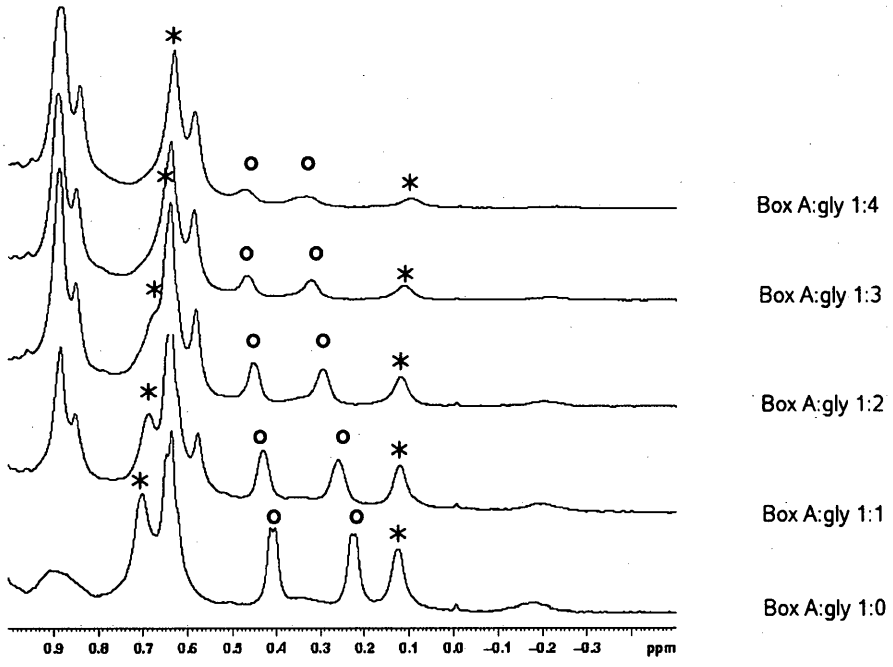


Fig. 3.7 Upfield region of ^{1}H monodimensional spectra of Box A during the titration with glycyrrhizin (0, 1, 2, 3, 4 equivalents of ligand). The resonances corresponding to the methyls of Val19 (*) and Val35 (°) are marked.

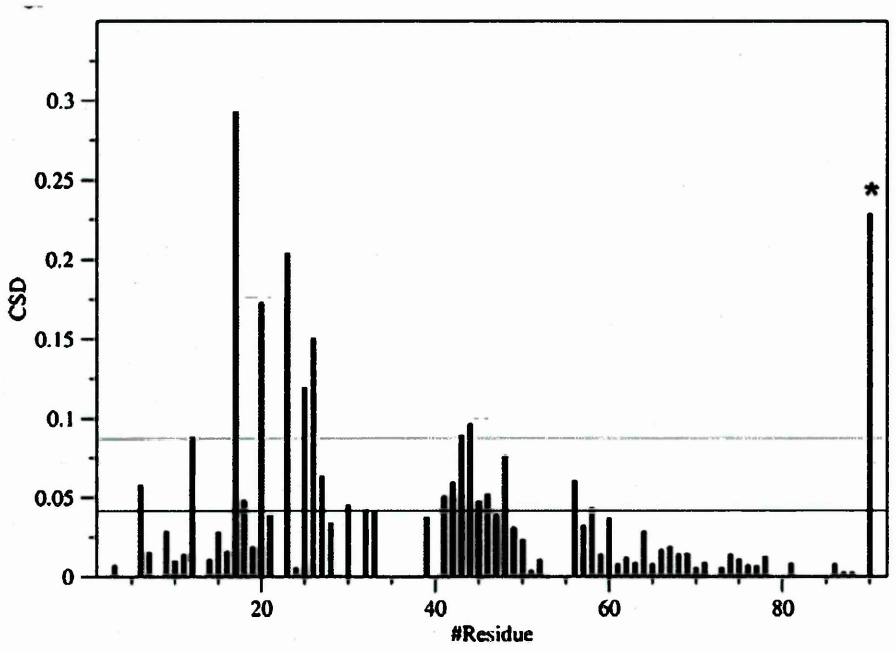


Fig. 3.8 Histogram representation of chemical shift differences (CSD) between CBNX bound (protein–ligand ratio 1:4) and free state of Box A. Secondary structure elements are reported on the top of the images. Gln20 sidechain is marked by an asterisk and reported with an increased order number. CSDs have been computed as reported in Materials and Methods section.

3.2.4 Influence of the binding on the protein diffusional dynamics.

T_1, T_2 and heteronuclear NOE values for free and bound Box A and for free and bound Box B are reported respectively in Fig. 3.9 and in Fig. 3.10. Average values of T_1, T_2 and heteronuclear NOE have been computed for helical regions and are reported in Table 3.1, together with correlation times τ_c that have been obtained from T_1 and T_2 using the empirical relationship of equation *2.3* (see Materials and Methods). In the table the residues that undergo large, slow and anisotropic motions are reported, together with the ones that are very flexible in the ns timescale and display a low (less than 0.6) value of heteronuclear NOE.

τ_c values for Box A and Box B reported in table 3.1 are very similar as expected

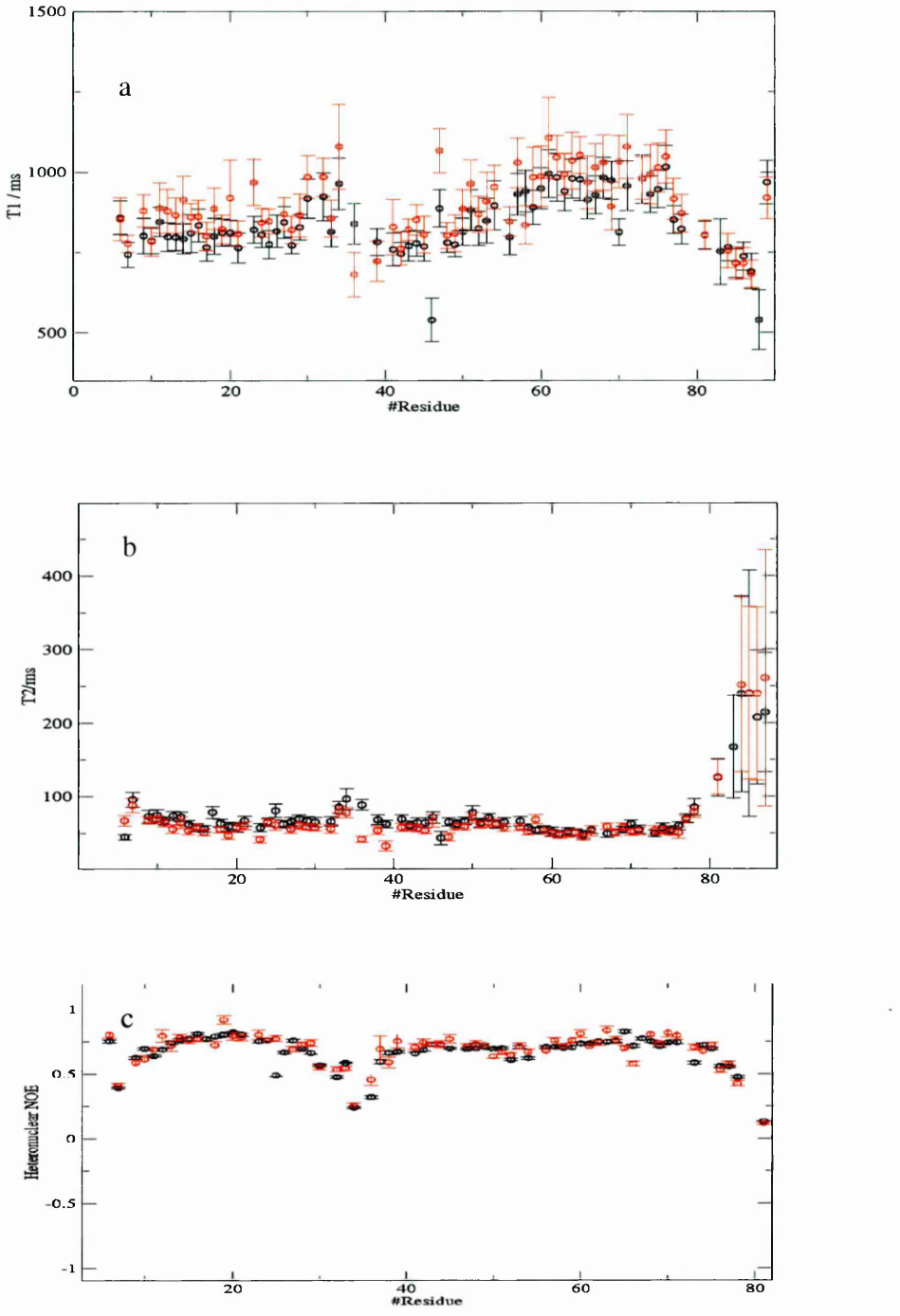


Fig. 3.9 T_1(a), T_2 (b) and heteronuclear NOE (c) of Box A (black) and of Box A–GL complex (red) .

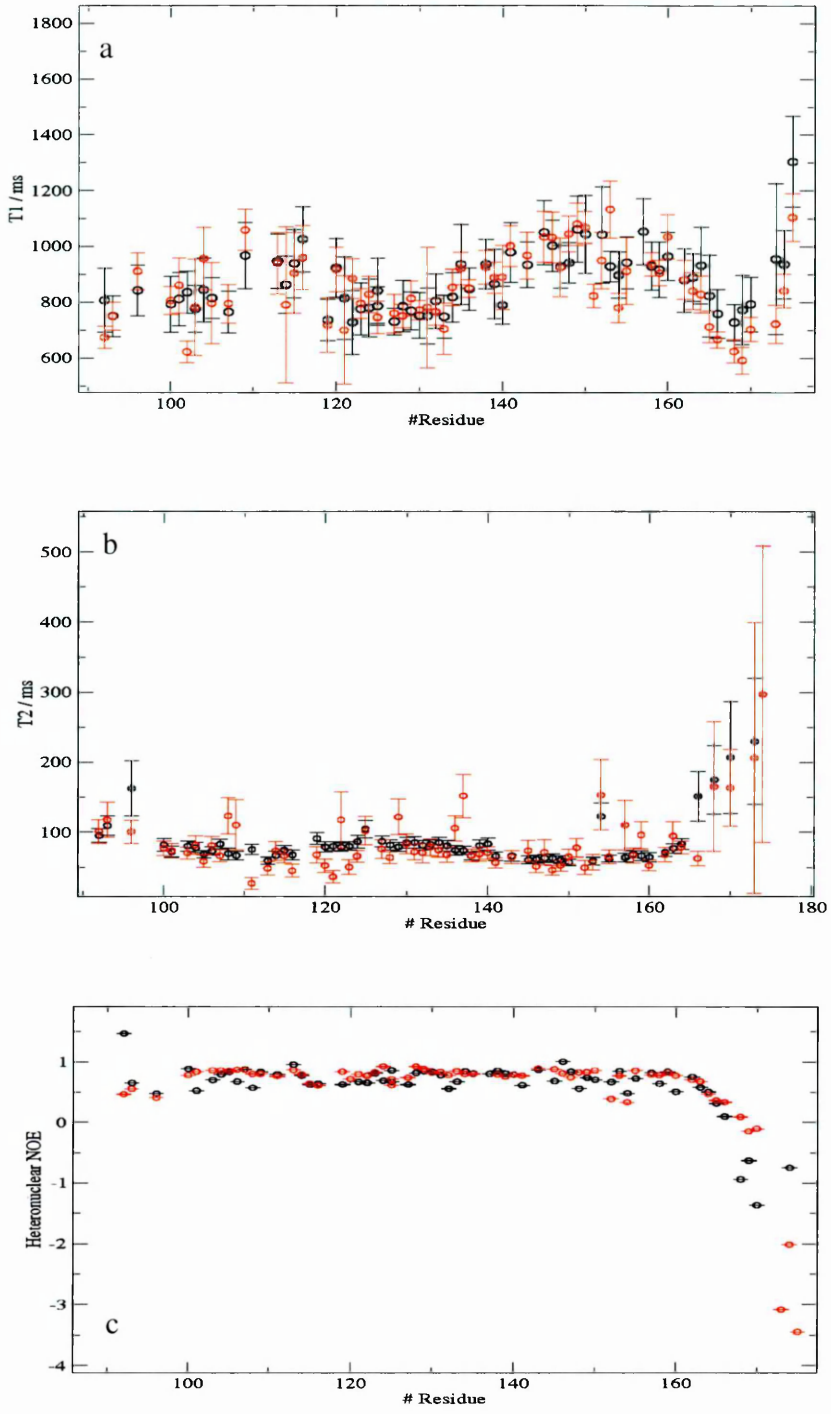


Fig. 3.10 T_1(a), T_2 (b) and heteronuclear NOE (c) of Box B (black) and of Box B–GL complex (red) .

System	H[a]	T_1[b]	T_2[b]	hetNOE	τ_c[b]	Excl. hetNOE [c]	Excl. T_2
BoxA	H1	820 ± 80	68 ± 9	0.71 ± 0.10	10.8 ± 0.2	7, 9, 11, 20, 25 30-36, 46, 52, 54, 73-89	none
	H2	811 ± 72	65 ± 7	0.68 ± 0.09			
	H3	934 ± 70	58 ± 6	0.71 ± 0.10			
BoxB	H1	899 ± 101	75 ± 10	0.72 ± 0.11	10.0 ± 0.1	108, 115, 119, 127, 132, 136, 141, 148, 158-175	148
	H2	805 ± 81	84 ± 10	0.73 ± 0.08			
	H3	953 ± 75	69 ± 12	0.71 ± 0.15			
BoxA+ GL	H1	909 ± 74	53 ± 4	0.73 ± 0.05	11.7 ± 0.5	7, 9, 10, 30-36, 50, 52, 77-89	23,47,61,64
	H2	905 ± 140	48 ± 10	0.69 ± 0.10			
	H3	969 ± 54	54 ± 4	0.71 ± 0.18			
BoxB+ GL	H1	892 ± 120	65 ± 29	0.76 ± 0.08	10.3 ± 0.4	115, 116, 137, 151-154, 164-175	none
	H2	810 ± 110	86 ± 31	0.79 ± 0.11			
	H3	967 ± 56	76 ± 29	0.69 ± 0.13			

a. Box A: Helix1, 13-27; Helix2, 39-52; Helix3, 55-72.
Box B: Helix1, 99-117; Helix2, 120-135; Helix3, 137-158.
b. T_1 and T_2 are expressed in ms; τ_c is expressed in ns .
c. Residues whose heteronuclear NOE have been excluded. Exclusion principles are explained in detail in the Materials and Methods section.

Table 3.1 T_1, T_2 and heteronuclear NOE values, correlation times and residues not included in the data analysis (*i.e.*, excluded residues) for the extraction of diffusive models due to excessive large motions.

from the overall similarity of the domains shape and fold, *i.e.* 10.8 ns for Box A and 10 ns for Box B respectively. These values are in good agreement with the ones previously determined (Laue et al 1993). The values of τ_c for the complexes are very similar to the ones of the free proteins, i.e. 11.7 ns and 10.3 ns for Box A-GL and Box B-GL respectively. The variation of τ_c is at least 10% of the free protein value, as expected

from the small size of the ligand that does not alter significantly the hydrodynamic radius of the protein. These values also indicate that the protein is still monomeric in solution after complexation.

The three main diffusional axis in Cartesian space D_x, D_y and D_Z are reported in Fig. 3.11, and their values are reported in Table 3.2. The formation of the complex between Box A and GL alters the overall shape of the diffusion tensor with respect to free protein (see Table 3.2): despite the persistence of an oblate character, it assumes a more spherical shape (Fig. 3.11b) due to the very close values of the three axes. Box B – GL complex diffusion tensor does not differ significantly from the one of the free Box B: despite a formal change in the symmetry of the ellipsoid, the values of D_x and D_y for the complex are similar and are also very close to the ones of the free protein within fitting errors.

3.2.5 Changes of internal Box A backbone mobility upon binding is monitored by ^{15}N spin relaxation

Relaxation data of both free and bound Box A display homogeneity in the values of T_2 ($<T_2>$ ~ 62 ±10 ms) and of heteronuclear NOE (<hetNOE> ~ 0.7 ± 0.1) excluding flexible regions, indicating that the secondary structure elements of the protein are not heavily perturbed by the binding of GL.

Low values of heteronuclear NOE (< 0.5) of the loop comprised between Helix 1 and Helix 2 indicate high mobility in the nanosecond timescale for these regions. The comparison of the differences between heteronuclear NOE datasets of the protein and of the complex reveals that the local dynamics for the free and bound protein in the ns timescale are very similar, i.e. differences between values for single residues are never larger than ±0.1. Also S2 values (Fig. 3.12a) extracted by means of the Lipari-Szabo formalism reveal this aspect: the differences of S2 between Box A-GL and Box A (reported in Fig. 3.13) are never higher than ±0.1 revealing no changes in local fast motions upon binding. Moreover, no relevant trends of S^2 or heteronuclear NOE variations can be detected as a function of secondary structure elements.

System	Model[a]	α^b	β^b	γ^b	D_x^c	D_y^c	D_z^c
BoxA	Asymm.	56 ± 21	43 ± 5	-40 ± 5	0.10 ± 0.01	0.13 ± 0.01	0.23 ± 0.02
BoxB	Ax.Simm.	-88 ± 6	-24 ± 6	-24 ± 6	0.13 ± 0.00	0.13 ± 0.00	0.23 ± 0.01
BoxA+GL	Asymm.	31 ± 37	71 ± 28	-66 ± 20	0.12 ± 0.07	0.13 ± 0.01	0.16 ± 0.01
BoxB+GL	Asymm.	-83 ± 32	-56 ± 3	-32 ± 4	0.12 ± 0.08	0.13 ± 0.06	0.26 ± 0.01

a. In the axially symmetric model $Dz \equiv D_{||}$, $D_x = D_y \equiv D_r$, $\beta = \gamma \equiv \phi$, $\alpha \equiv \theta$
b. Degrees
c. 100 X MHz

Table 3.2 Diffusive models and relative angular parameters and axis magnitude

T_2 values for the bound protein are lower than in the free form (with the exception of C and N termini) in the region comprising residues 10 and 54: in particular, this aspect can be observed for residues ranging from 29 to 40 that comprise the N terminal part of Helix 1, the C terminal part of Helix 2 and a flexible loop (Fig. 3.9). This region of the protein roughly corresponds to the set of residues whose chemical shift is mostly perturbed upon binding. T_2 variations are reported in Fig. 3.14: the change of values of T_2 can be usually related to chemical exchange, hence this graph suggests that the relaxation of ^{15}N spin in this region reflects the binding kinetics of the ligand. In this sense, T_2 values of Helices 1 and 2 are perturbed by the presence of the ligands: the greatest variations of T_2 are reported in the loop region, suggesting that its slow motion is enhanced by the binding. This result is compatible with the absence of great variations in heteronuclear NOE between the complexed form and the free form of Box A, providing an indication that only slow motions (milliseconds time scale) are influenced by the binding.

K_{ex} parameters have been extracted using the Lipari-Szabo model and they are

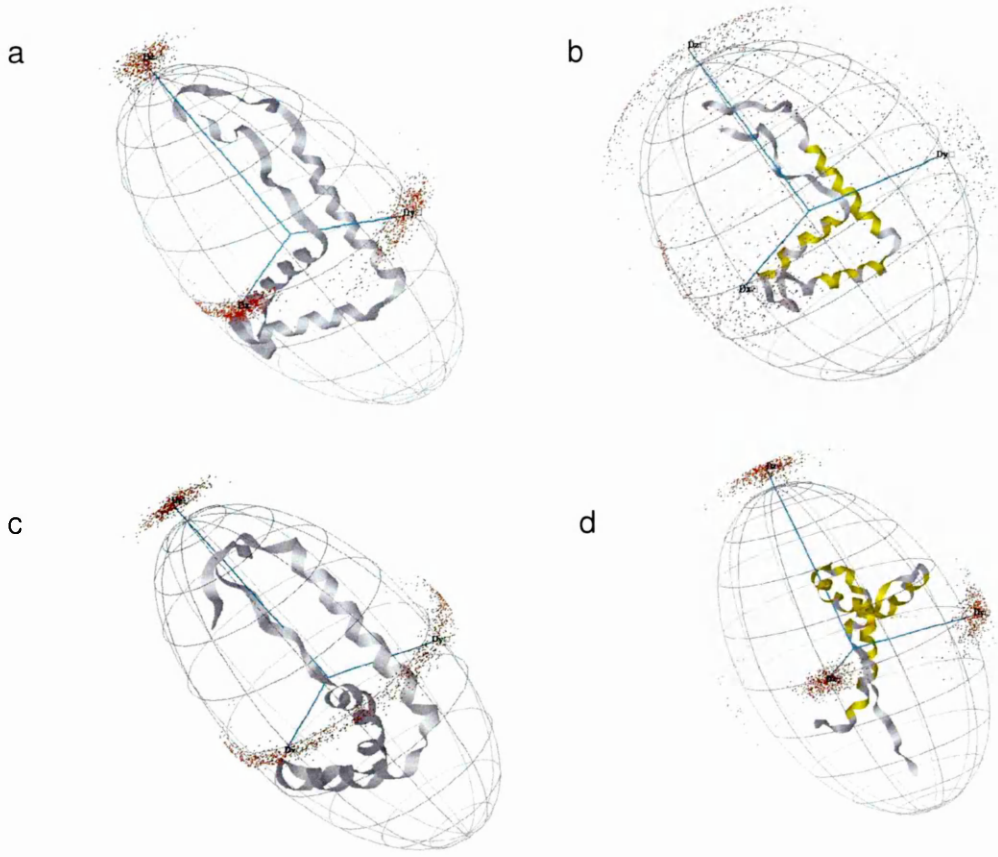


Fig. 3.11 Ellipsoidal representation of diffusion tensors computed from T_1 and T_2 data for Box A (a), Box A–GL (b), Box B (c) and Box B–GL (d). The center of the axes is located in the centre of mass of the protein. The dots reported on the surface of the ellipsoids are representative for the error distribution of the D_z, D_y and D_x representation computed according to the statistical procedure described in the Materials and Methods section.

reported in Fig. 3.12b. The free protein displays K_{ex} components that are quite spread over the whole structure, whereas upon binding the K_{ex} components are located on Helices 1 and 2 and they identify the same binding site previously identified though chemical shift mapping. The presence of exchange contributions is also evidenced by the increased linewidths of peaks corresponding to methyl groups of residues 15 (Helix 1) and 35 (Helix 2) in the 1D spectrum reported in Fig. 3.7. Residues 23 and 47 exhibit an anomalous T_2 (i.e., too low to exclude anisotropic motions contributions, see Table 3.1) that indicates the presence of anisotropic motion that can lead to an overestimation of

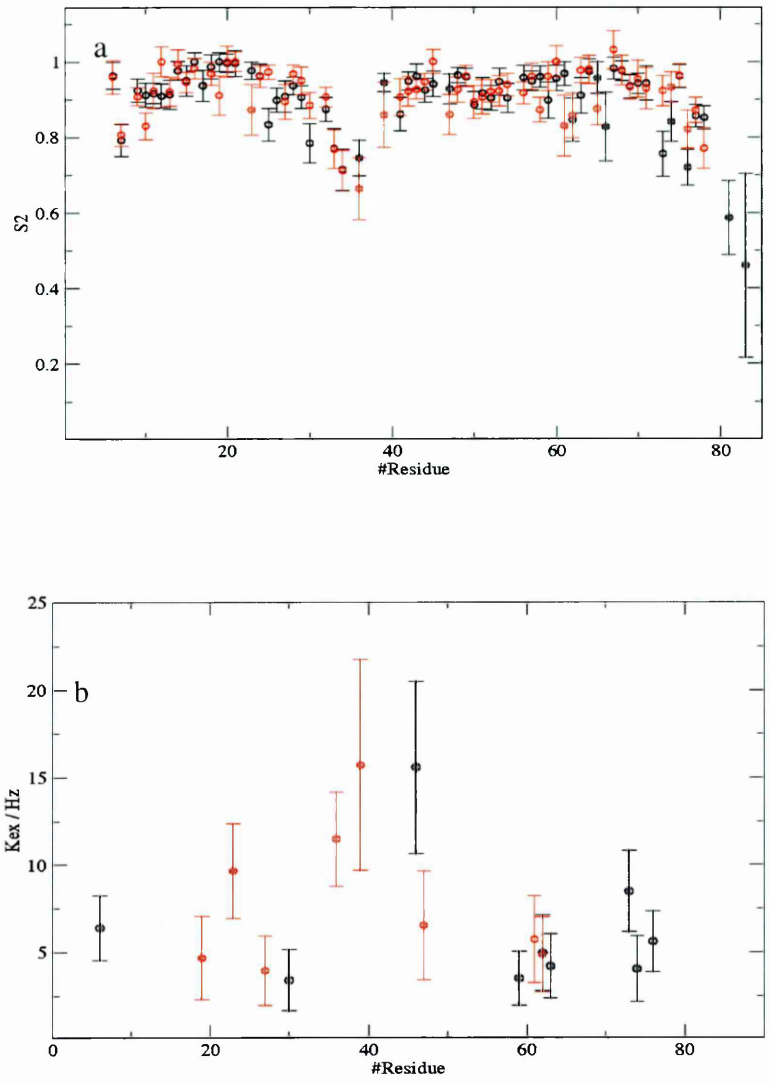


Fig. 3.12 Lipari Szabo S^2 (a) and K_{ex} (b) parameters extracted from the analysis of relaxation data of free (black) and GL bound (red) Box A.

the proper exchange contribution. Residues 61 and 62, located on Helix 3, exhibit an exchange component: this can be partially explained by the contact with the residues belonging to GL binding site and located on Helix 1 and Helix 2.

Interestingly, the free Box A dynamics reveals the presence of exchange components in the C terminal region which could be related to an unstable helical

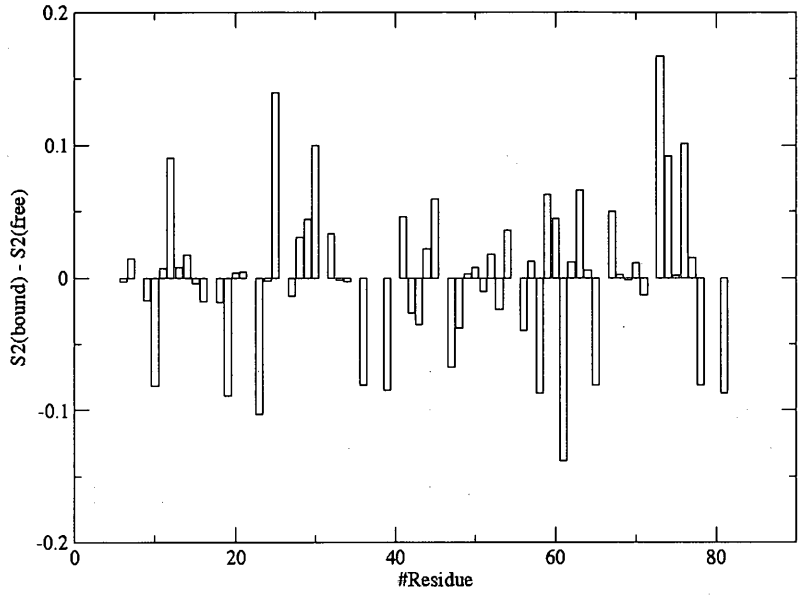


Fig. 3.13 Histogram representation of the difference between S^2 values of the bound and the free form of Box A.

folding of the residues between 68 and 75. This is in good agreement with REMD calculations (see Materials and Methods section) performed on free proteins that for this region display a quite high Cα RMSF (Fig. 3.15b). This information complements the description of a local helical fold in this region suggested by the deposited PDB structure and is in good agreement with the Cα-RMSF computed on the bundle of structures of Box A (see Fig. 3.15a). Taken together, these data suggest that, with respect to the remaining part of the protein, in this region the helical motifs breath between different conformations due to a decreased structural stability. The timescale of REMD simulations is roughly 80 ns distributed over all the replicas: the agreement between this timescale and the one (*i.e.*, ms) represented by a bundle of structure is quite surprising, but it can indicate that the essential motions that govern exchange contributions are well represented also by motions that take place in the high

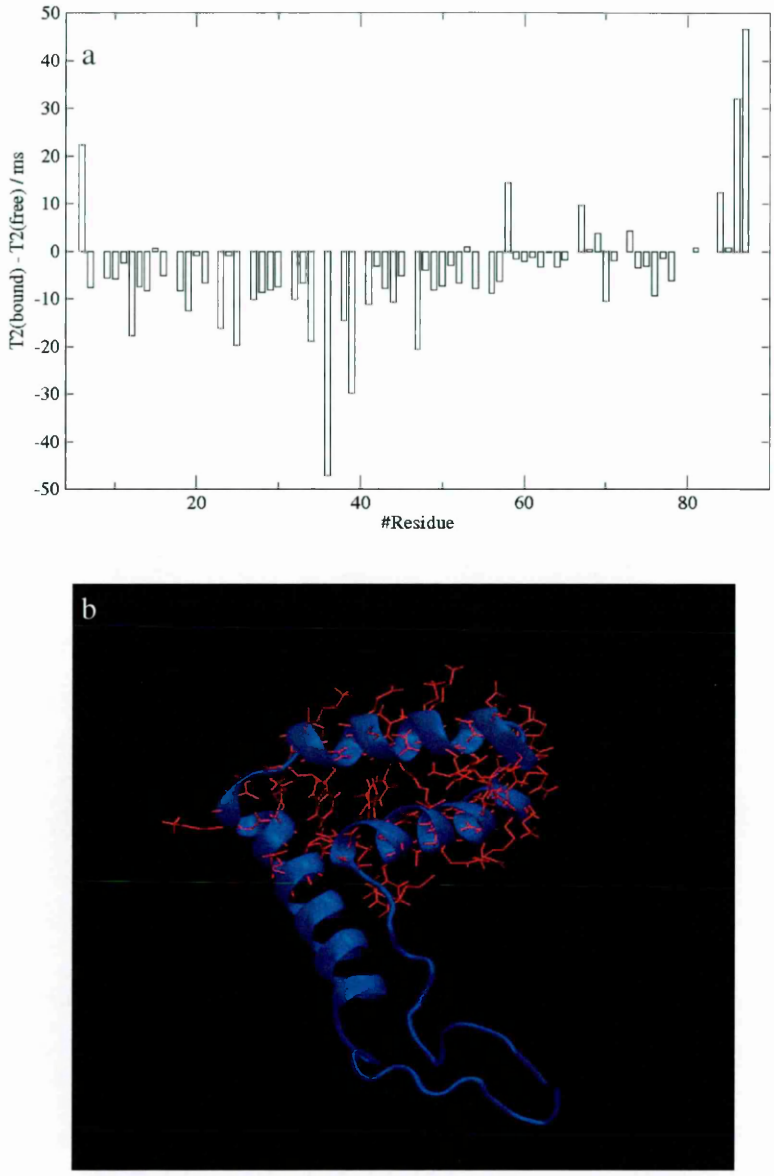


Fig. 3.14 Histogram representation of the difference between T_2 values of the bound and the free form of Box A. (a). The residues are reported on the structure in red (b).

variation of S2 (> 0.2), confirming that for Box B motions in the ns timescale are perturbed by the binding of GL.

nanosecond timescale and that well represent the overall correlation time of the protein.

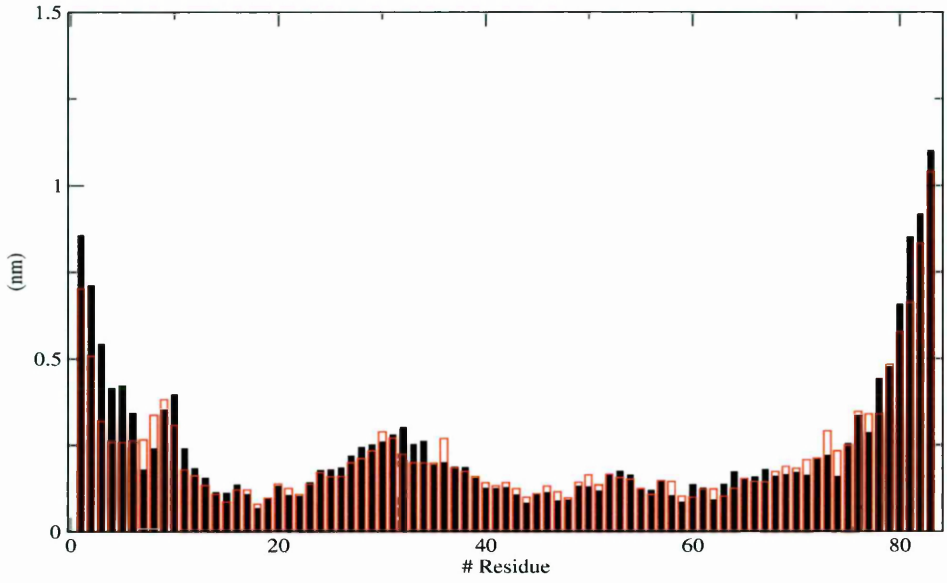


Fig. 3.15 Cα RMSF of free Box A computed on the deposited bundle of structures 1AAB (black) and from a REMD simulation performed as described in Materials and Methods section (red).

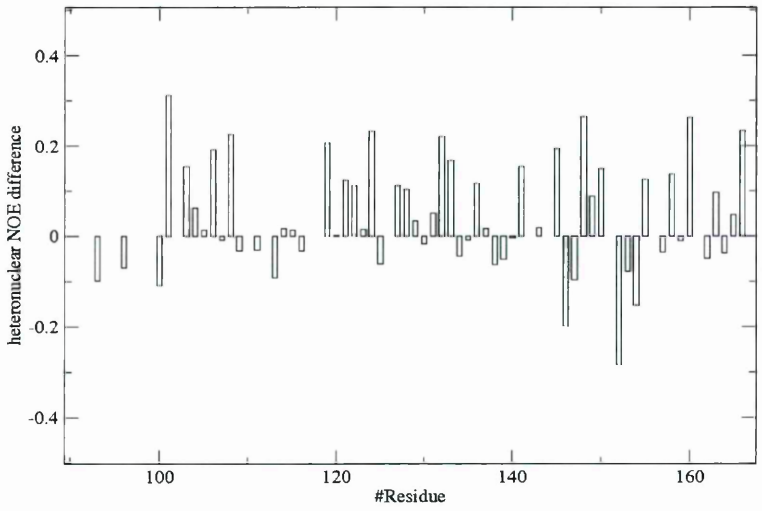


Fig. 3.16 Histogram representation of the difference between *heteronuclear NOE* values of the bound and the free form of Box B.

3.2.6 Changes of internal Box B backbone mobility upon binding is monitored by ^{15}N spin relaxation

As shown in Fig. 3.9c and 3.10c, average heteronuclear NOE values of Box B and of Box A are very similar (~ 0.7 , see Table 3.1). Upon binding, as reported in the histogram of Fig. 3.16, all the values change and the variation in many cases is quite high, indicating that the local dynamics sampled on the nanosecond is changed by the binding The average variation of hetNOE values is positive: this suggests that GL rigidifies the fast (ps-ns) motions of Box B upon binding. This result is also partially reflected in the DS2 values displayed in Fig. 3.18 (3.17 reports the S2 and Kex values obtained by Lipari-Szabo analysis of relaxation data): the plot reported suggests that residues 101, 109 and 122 are the ones whose ns dynamics is mostly influenced by the binding and that loose their conformational entropy.

T_2 values for the free Box B are almost constant ($T_{2,\,avg}$ ~ 70ms) in the structured regions, with the second helix having T_2 values slightly higher than the ones for the Helix 1 and 3. Upon binding, the T_2 values mostly perturbed are the ones of the residues between 105 and 138, that corresponds to half of the Helix 1 and the complete Helix 2 (see Fig. 3.17c), that fits well with the binding site for GL identified by chemical shift mapping. Similarly to Box A, the T_2 values are lowered on average and there are some outliers with respect to the overall trend (i.e., T_2 higher than the average) mainly located on Helix 2. K_{ex} obtained for the free and bound state of Box B are reported in Fig. 3.17. The free protein reveals, like Box A, a few residues that could be fitted according to the Lipari-Szabo model using a K_{ex} component. These residues are spread over the entire sequence: when GL complexes Box B exchange components can be identified only on the Helices 1 and 2 and they are not present on Helix 3. Noteworthy, the slowest term is associated to Phe101: this observation is in good agreement with the direct contact with GL evidenced by 2D X-filtered NOESY experiments (see further). Moreover, residues 123 and 125 are among the ones whose chemical shift is displaced mostly by ligand binding. Taken together these three residues identify the site that displays the greatest variation of S^2 (> 0.2), confirming that for Box B motions in the ns timescale are perturbed by the binding of GL.

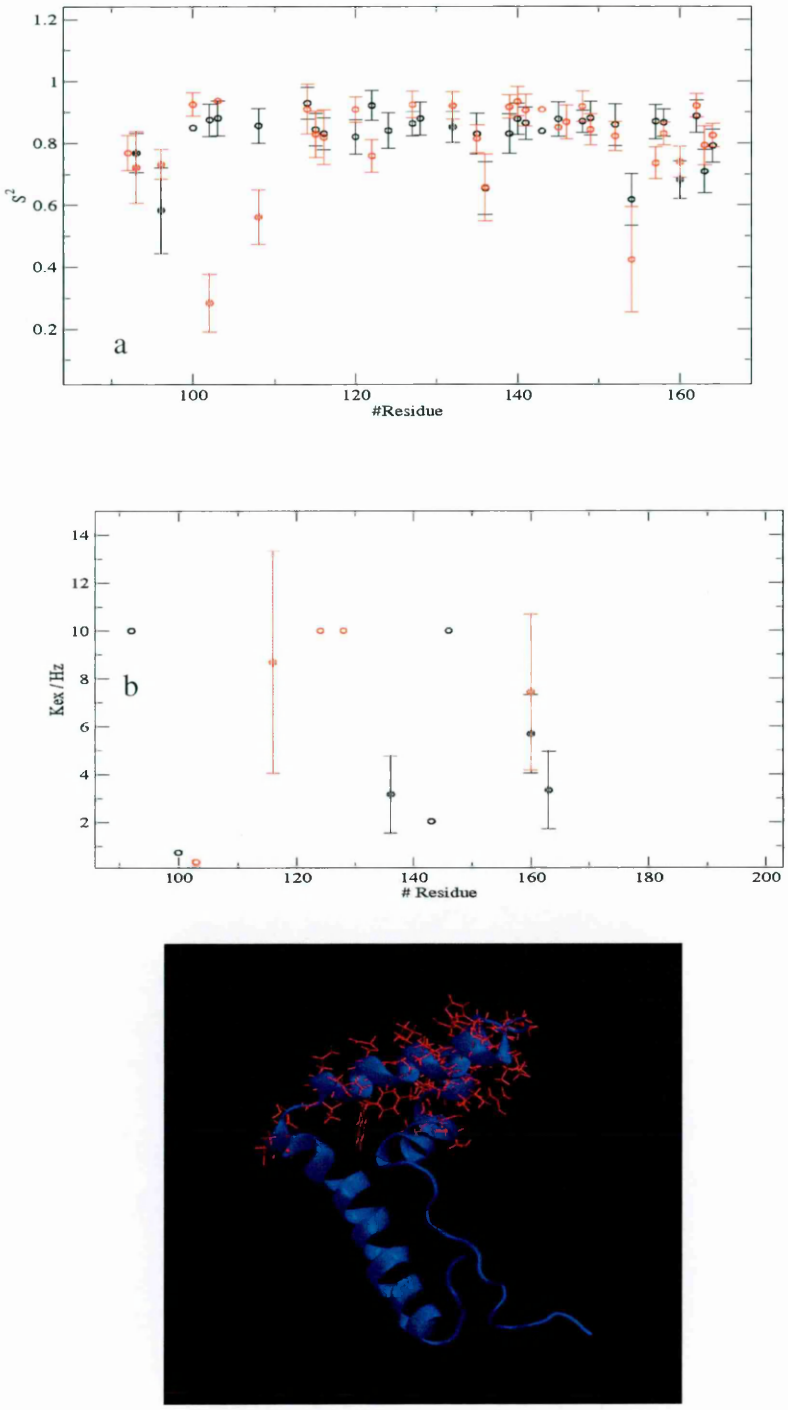


Fig. 3.17 Lipari Szabo S2 (a) and Kex (b) parameters extracted from the analysis of relaxation data of free (black) and GL bound (red) Box B. The residues mostly perturbed according to T2 measurements are reported in red on Box B structure (c).

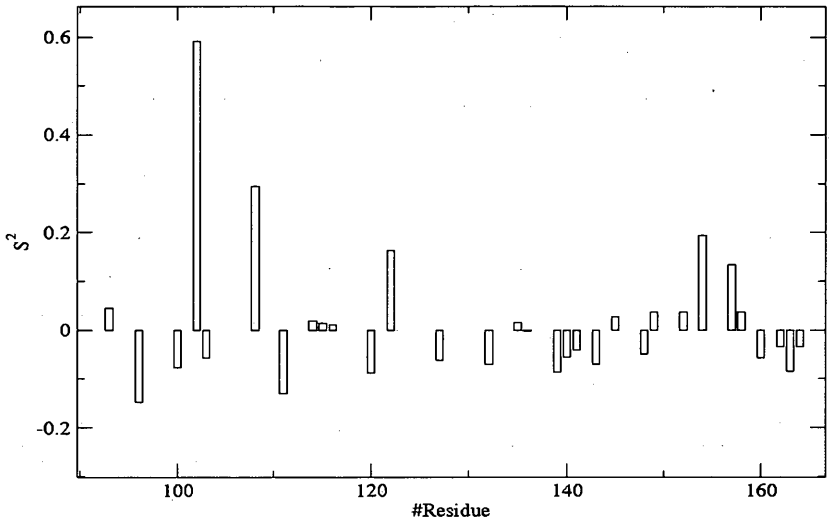


Fig. 3.18 Histogram representation of the difference between S^2 values of the bound and the free form of Box B.

3.2.7 Determination of K_d by NMR and fluorescence

Dissociation constants were estimated by monitoring the change of individual resonance chemical shifts of the single boxes as a function of ligand concentration. Plots of the chemical-shift changes of two representative residues for box A (R23) and box B (R109) as a function of ligand concentration are shown in Fig. 3.19 Assuming a simple binary interaction between single boxes and GL, we derived K_d values of 103 ± 25 μM and 87 ± 35 μM for box A and box B, respectively. Because of the low quality of the NMR spectra of the full-length protein in the presence of GL at stoichiometric ratios higher than 1:4 (HMGB1:GL), we could not follow the NMR signals up to saturation to obtain a reliable estimate of the binding constants. We investigated whether intrinsic fluorescence could be used to monitor the interaction of HMGB1 with GL (Fig. 3.20). Two conserved tryptophans (Trp48 and Trp132 in box A and box B, respectively) are buried inside the hydrophobic core of the domain and do not belong to the proposed ligand binding site; as expected, no noticeable change in their fluorescence was observed (data not shown). Two tyrosine (Tyr15 and Tyr108 in Box A and Box B, respectively)

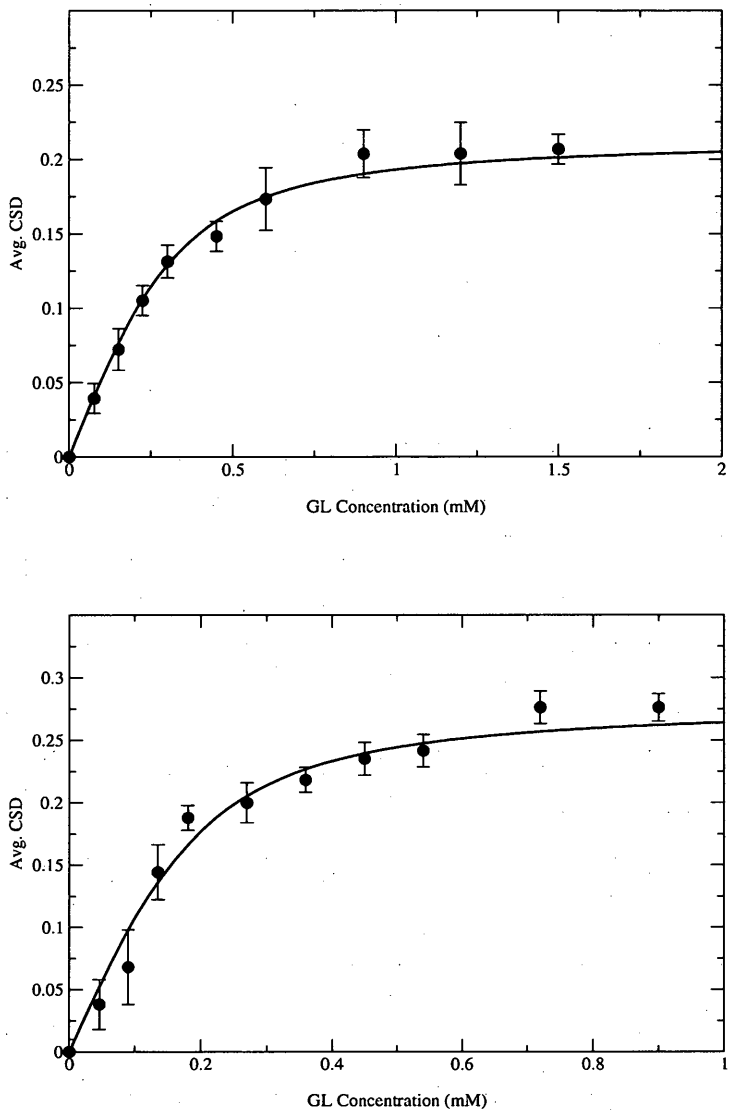


Fig. 3.19 Weighted average of the Arg23 (a) and Arg109 (b) amide ^{1}H and ^{15}N chemical-shift changes as a function of added glycyrrhizin values have been computed as reported in Materials and Methods. A simple binary interaction has been assumed between single boxes and glycyrrhizin.

are close to the residues showing chemical-shift perturbations; tyrosine fluorescence decreases upon GL addition (3.20). Nonlinear least-squares fitting of the data yields dissociation constants of 170 ± 3 μM for box A and 140 ± 3 μM for box B, assuming a 1:1 stoichiometry of isolated boxes and GL. The binding of GL to both HMG boxes,

together with the structural independence of the boxes inside HMGB1, strongly suggest that the full-length protein has two independent binding sites for GL. We therefore fitted

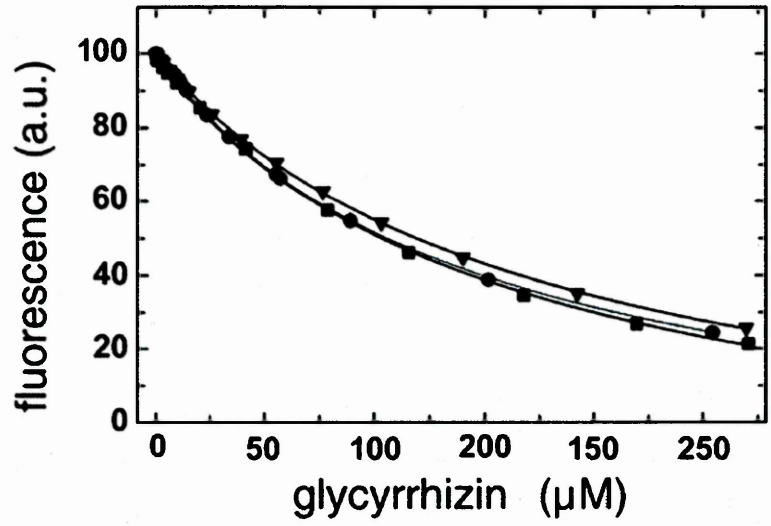


Fig. 3.20 Tyrosine fluorescence (excitation 273 nm, emission 325 nm) was measured (in arbitrary units; a.u.) upon titration of glycyrrhizin into solutions of Box A (triangles), Box B (circles), and FL-HMGB1 (squares). Symbols correspond to experimental data, whereas continuous curves are derived according to equations described in the Materials and Methods section.

fluorescence data for the full-length protein assuming a 1:2 (HMGB1:GL) binding stoichiometry and two independent dissociation constants, KdA and KdB, for the binding site on box A and box B, respectively. Nonlinear least-squares fitting of the data did not yield two statistically different values for KdA and KdB but rather resulted in a common value of 156 ± 3 μM. This value is similar to the value obtained for the single boxes, further suggesting that the binding of GL to the full-length protein is noncooperative and does not result in an improved binding affinity relative to single boxes. These results also suggest that the acidic tail does not interfere in the binding of GL to HMGB1.

4. Structural models for the complexes

4.1 Summary

In the previous chapter binding of CBNX to Box A and of GL to Box A and to Box B was experimentally probed by chemical shift perturbation methods and by fluorescence spectroscopy: this procedure allowed a clear identification of the ligand binding sites on the surface of the protein. In the present chapter the problem of providing an atomic level description of the binding modes of CBNX and GL to the HMGB1 boxes is faced. Due to the quite high K_d values obtained for the complexes, protein-ligand NOE peaks obtained were insufficient for a complete structure determination performed adopting a standard approach (*i.e.* NOE derived distance based methods). This fact led us to adopt a combined computational and experimental approach for protein-ligand complex structure determination.

A hybrid quantum chemistry and molecular dynamics (MD) based method with back calculation and *a posteriori* confrontation with experimental data was adopted for CBNX: a density functional theory (DFT) approach was used for refining the ligand charges derived from a standard molecular mechanics force field and the whole complex dynamics was simulated using MD protocols.

A data driven docking procedure was instead adopted for GL, using all the available experimental data as restraints. This choice was driven by the poor reliability of available protein force fields for the treatment of sugar units in GL: even a quantum mechanics based strategy like the one adopted for CBNX would not have led to a reliable set of charges for sugar moieties, *i.e.* starting from a very approximate set of charges. Hence, using the whole set of data available from NMR experiments in order to get intermolecular restraints for protein-ligand docking, despite the great problems faced by this kind of algorithm, offered the more accurate and fastest solution available to the problem of determining the structure of the Box A-GL and of the Box B–GL complexes.

4.2 Results

4.2.1 Experimental measurements on Box A–CBNX complex

STD measurements have been performed on all the systems previously described (see previous chapter), suggesting the proximity of the ligand methyls 12, 19, 24, 25, 26 and 28 (see Materials and Methods, Fig. 1.1) to the surface of the protein in all the cases considered, confirming the importance of the triterpenic scaffold in the binding, as previously demonstrated. STD measurements performed on Box A–CBNX complex are reported in Fig. 4.1. This kind of experiment is usually performed with success on complexes formed by large proteins (>10kD) and flexible ligands, because this experimental condition ensures the optimal magnetization transfer from protein to ligand. In this way it is possible, in principle, to quantify the distance of ligand protons from the surface of the protein and to evaluate the kinetics of binding through the determination of k_{on} and k_{off} for single protons (Mayer and Meyer, 2000). In this case the rigidity of the triterpenic scaffold could bias the relative amount of transfer magnetization and consequently bias any quantitative evaluation of distances. However, it was possible to identify with good confidence the CBNX methyl groups that are in contact with the protein surface hence defining well the ligand-protein interaction surface.

In order to establish the protein-ligand contacts, X filtered NOESY measurements were performed (Breeze, 2003) adopting the ω_2 $^{15}N^{13}C$ filtering scheme (Fig. 4.2) in order to identify the hydrogen atoms of the ligand along the direct dimension and the one of protein sidechains along the indirect one: protein signals were previously identified by means of and ^{15}N edited HSQC-NOESY. In Fig.4.2 a portion of the X-filtered spectrum obtained for the complex Box A–CBNX is reported: peaks inside the box correspond to contacts between methyl and methylene GL protons and protons of protein aromatic sidechains of residues Tyr15 and Phe37. This experiment allowed the identification of up to 10-15 intermolecular contacts, but the signal intensity and broadening (that depended presumably on different positions assumed by the ligand in the protein binding site) disallowed any precise evaluation of inter proton distances. Only qualitative evaluation on the distances involved was done and classified according

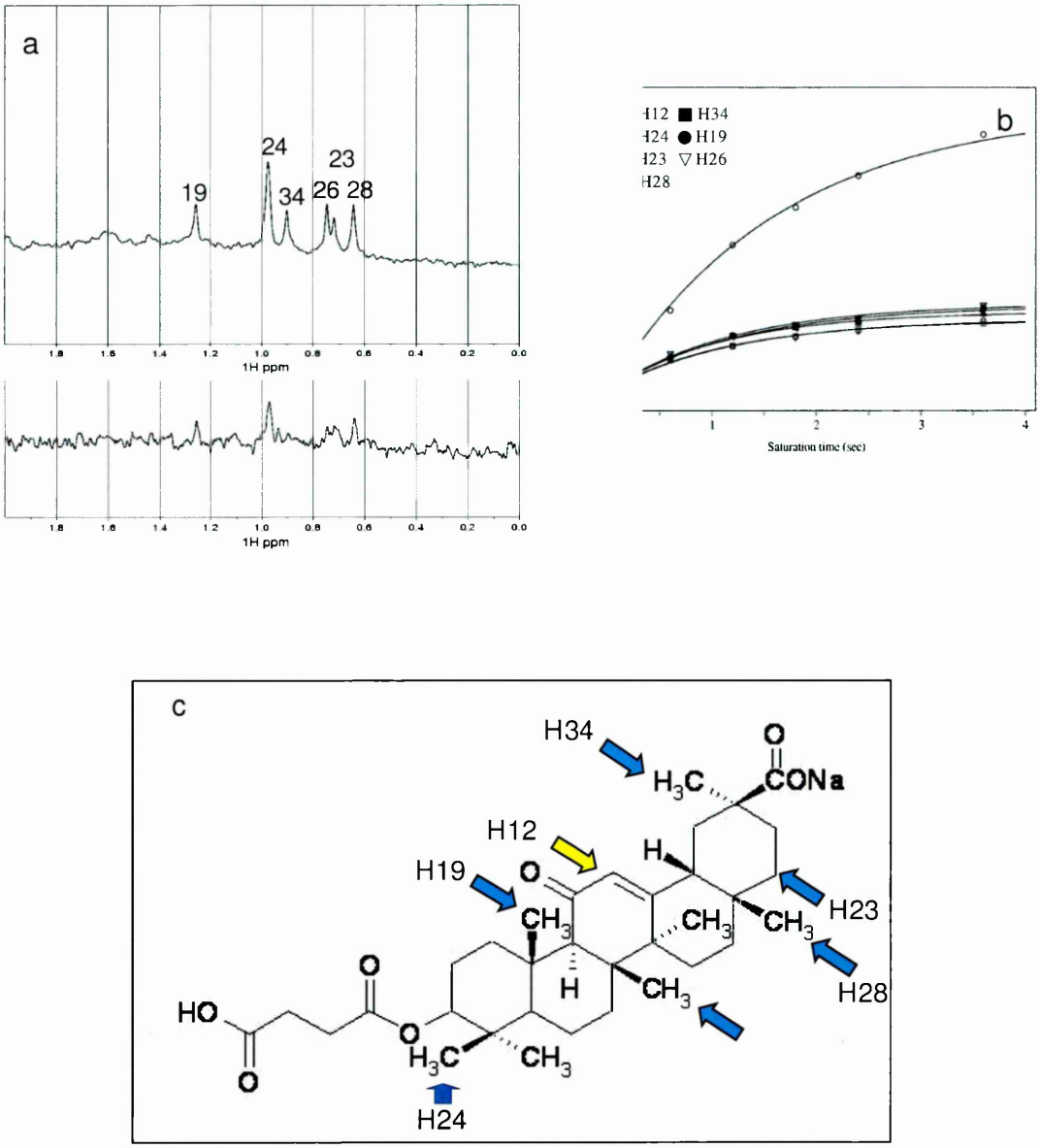


Fig. 4.1 CBNX protons in proximity of Box A surface in the Box A–CBNX complex detected by STD experiments: a) aliphatic region of the STD spectra of the complex recorded with a saturation time of 0.2 s (lower spectrum) and of 2.0 s (upper spectrum); b) time dependent saturation curves for detected protons; b) protons detected by STD measurements (in yellow: H12).

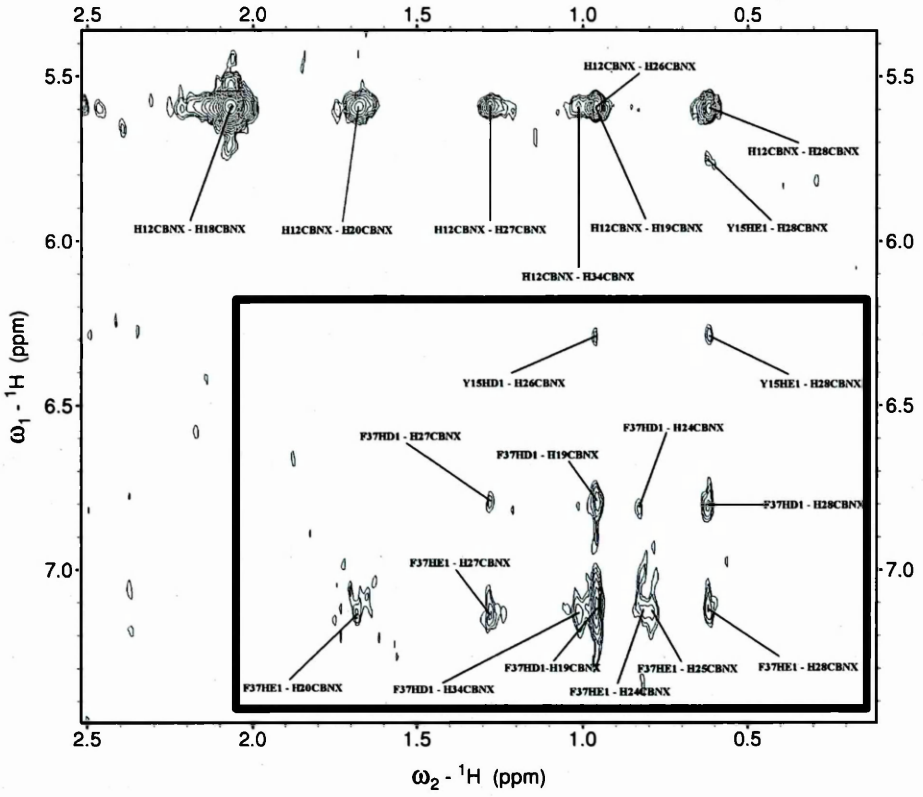


Fig. 4.2 600 MHz ω2 X-filtered NOESY spectrum portion. Intra- and Intermolecular (inside the square) NOE crosspeaks are indicated.

to the intensities of signals (see further). The contacts between Box A and CBNX are reported in Table 4.1.

4.2.2 QRFF–MD simulation of the BoxA – CBNX complex

Using all the available experimental information about the protein-ligand interaction surface and ligand orientation (*i.e.*, CSD, X-filtered NOESY and STD spectra) we attempted to place the ligand manually on the protein surface and also used AUTODOCK for an independent (unconstrained) docking. In both cases, we ended up with a configuration of the type illustrated in Fig. 4.3a in which the ligand is positioned on the surface of both Helix 1 and 2. This structure was equilibrated in explicit solvent

NOE Assignment	Signal Intensity	Distance (Å) QRFF	Distance (Å) Unrefined
Tyr15,Hδ – H26	weak	5.6 ± 0.4	3.9 ± 0.3
Tyr15,Hδ – H28	weak	3.4 ± 0.4	3.8 ± 0.5
Tyr15,Hε – H28	weak	3.0 ± 0.4	4.6 ± 0.6
Phe37,Hε – H20	weak	6.0 ± 0.5	12.6 ± 0.8
Phe37,Hδ – H24	weak	8.2 ± 0.6	3.2 ± 0.4
Phe37,Hε – H24	weak	6.5 ± 0.5	3.6 ± 0.5
Phe37,Hε – H25	weak	8.8 ± 0.4	5.4 ± 0.5
Phe37,Hδ – H26	strong	3.2 ± 0.4	7.4 ± 0.4
Phe37,Hε – H26	strong	2.9 ± 0.3	8.1 ± 0.9
Phe37,Hδ – H27	weak	6.9 ± 0.3	9.0 ± 0.4
Phe37,Hε – H27	medium	7.4 ± 0.4	10.0 ± 0.7
Phe37,Hδ – H28	medium	2.4 ± 0.3	11.3 ± 0.5
Phe37,Hε – H28	medium	3.7 ± 0.5	12.0 ± 0.9
Phe37,Hε – H34	weak	9.0 ± 1.1	14.9 ± 0.8

Table 4.1 Observed NOEs, features (*i.e.*, intensity) of the signals, and internuclear distances obtained from r^{-6} averaging over the trajectories of REMD simulations, based either on the unrefined force field or on the quantum-refined (QRFF) force field.

Complex	Ambiguous Restraints	Unambiguous Restraints
Box A-GL	17, 20, 23, 25, 26, 42, 43	15-H28, 37-H27, 37-H19
Box A-GL	107, 109, 113, 122, 123, 125, 129,133	102-H34, 102-H28, 101-H24

Table 4.2 List of ambiguous and unambiguous restraints used for docking calculations for Box A–GL and Box B–GL complexes. Residues whose chemical shift was affected by the binding are reported as ambiguous restraints. The unambiguous restraints reported refer to NOEs arising from protein-ligand contacts, and are reported in the form of a residue sidechain proton (see text)– ligand proton contact.

using the GROMOS 453a1 FF (see Materials and Methods section) for 1 ns (region I). Subsequently, the QRFF based MD was used (i.e., standard topology file was replaced

by the QRFF modified one) and the system was let free to evolve for 5 ns at room temperature. The time variation of the ligand–protein interaction energy is reported in Fig. 4.4 and clearly shows that after about 2 ns (region II) the system spontaneously found a more thermodynamically favored basin (DE = 75 ± 12 kJ/mol). The analysis of the binding energy of the complex shows that the gain corresponding to the structural transition (from energy basin I to II) is larger for the electrostatic component (DE = 50 ± 10 kJ/mol) than for the dispersive term (DE = 25 ± 5 kJ/mol). The separate energetic components are reported in Fig. 4.5.

Corresponding to the transition from the starting complex configuration to the final energetic equilibrium plateau (II), the complex underwent a substantial structural rearrangement, as reported in Fig. 4.3b, which illustrates the structure obtained from averaging over the configurations visited by the system during the final 2 ns of MD. At the transition (Fig. 4.4), the ligand reoriented itself almost perpendicularly to Helix 1, thus moving away from Helix 2: as expected, the structure of the ligand turned out to be rather rigid over this period, with a RMSD of 0.8 ± 0.2 Å (computed on the triterpenic scaffold of CBNX). This binding mode is consistent with chemical shift difference data upon binding (Fig. 3.8) due to the proximity of the ligand with the residues whose chemical shift was mostly perturbed by the binding.

The binding of CBNX and its rearrangement during the run do not alter the global conformation of the protein, as demonstrated by the relatively low Cα RMSD (2.2 ± 0.4 Å) computed on structured portion of the protein (*i.e.*, from residue 12 to residue 72) and by comparison of the Ramachandran plots of the two structure (Fig. 4.6) that ensures that no local structural distortions occurred during the ligand rearrangement on the surface of the protein. This finding is in total agreement with circular dichroism data (see Fig. 3.4). However, Helices orientation analysis reveals local modifications induced by the binding: the major change we observe pertains to the relative orientations of Helices 1 and 2 (Fig. 4.7), which change by up to ~ 20°. After the transition, the backbone remains rather rigid: the C_α RMSD of the backbone over the last 2 ns amounts to 1.1 ± 0.3 Å.

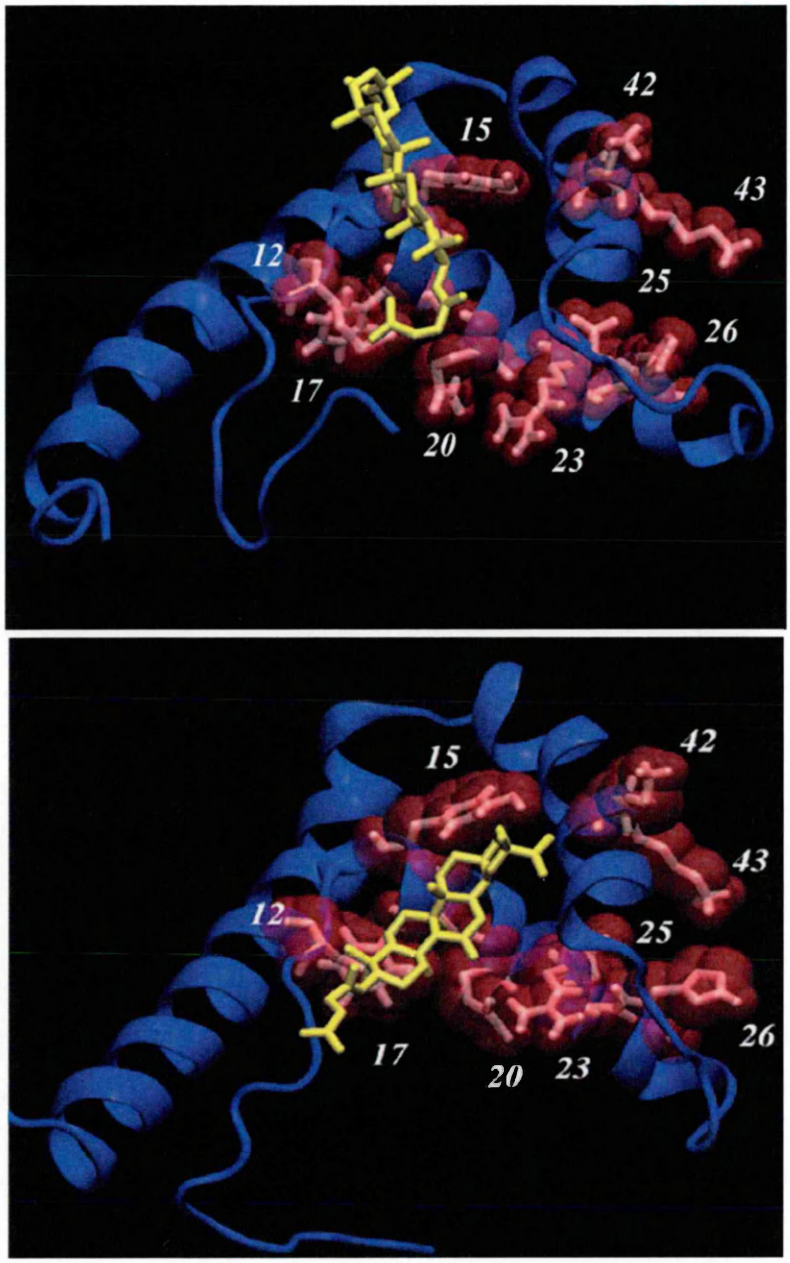


Fig. 4.3 Box A–CBNX complex structures at the beginning (a) and at the end of the 5 ns MD simulation. The structure reported in (a) is the one obtained through AutoDock docking calculations.

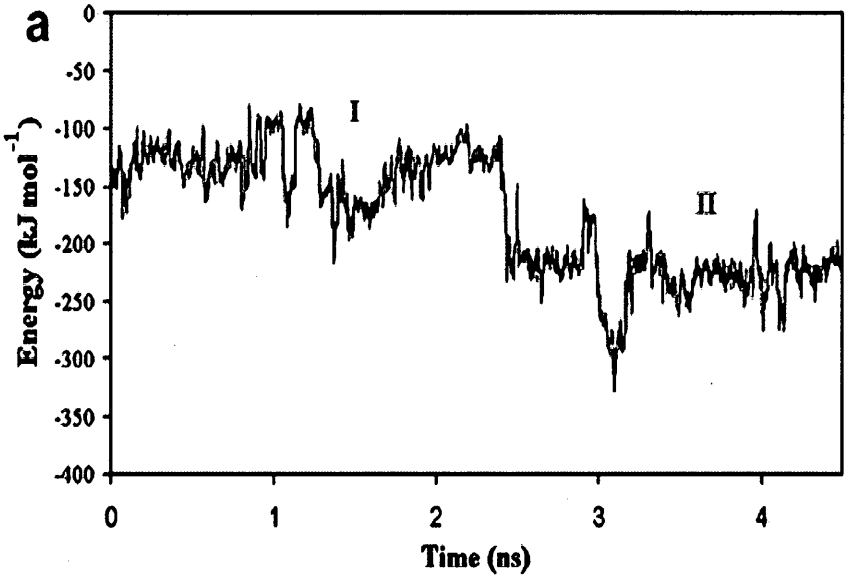


Fig. 4.4 Total interaction potential energy of the protein-ligand system.

Several residues are involved in a hydrophobic contact with CBNX, in particular, Arg9, Gly10, Lys11, Met12, Tyr15, Ala16, Phe37 and Ser41 The structure is furthermore stabilized by the formation of a single hydrogen bond between Ser38 sidechain hydroxyl and the carboxylic oxygen O34 of CBNX: this bond corresponds to a distance of 2.7 ± 0.5 Å but it has a lifetime of at most 50 ps (Fig. 4.8).

REMD simulations in NPT and NVT ensemble in the range 300-380K were performed in order to test the stability of the protein-ligand configuration found during the 5 ns MD simulation. The structure resulted to be very stable and a cluster analysis performed according to the minimum distance matrix method (Daura et al. 1999) revealed the presence of 22 ligand clusters with a cutoff of 0.1 nm that are shown in Fig. 4.9. Moreover, the REMD trajectories analysis confirmed the data obtained by MD simulations. In addition, the analysis of the NPT REMD trajectories for the ligand–protein interaction energy in the range 300–400 K reveals quite surprisingly that its contribution to the specific heat C_p does not depend on temperature (Fig. 4.10) and only

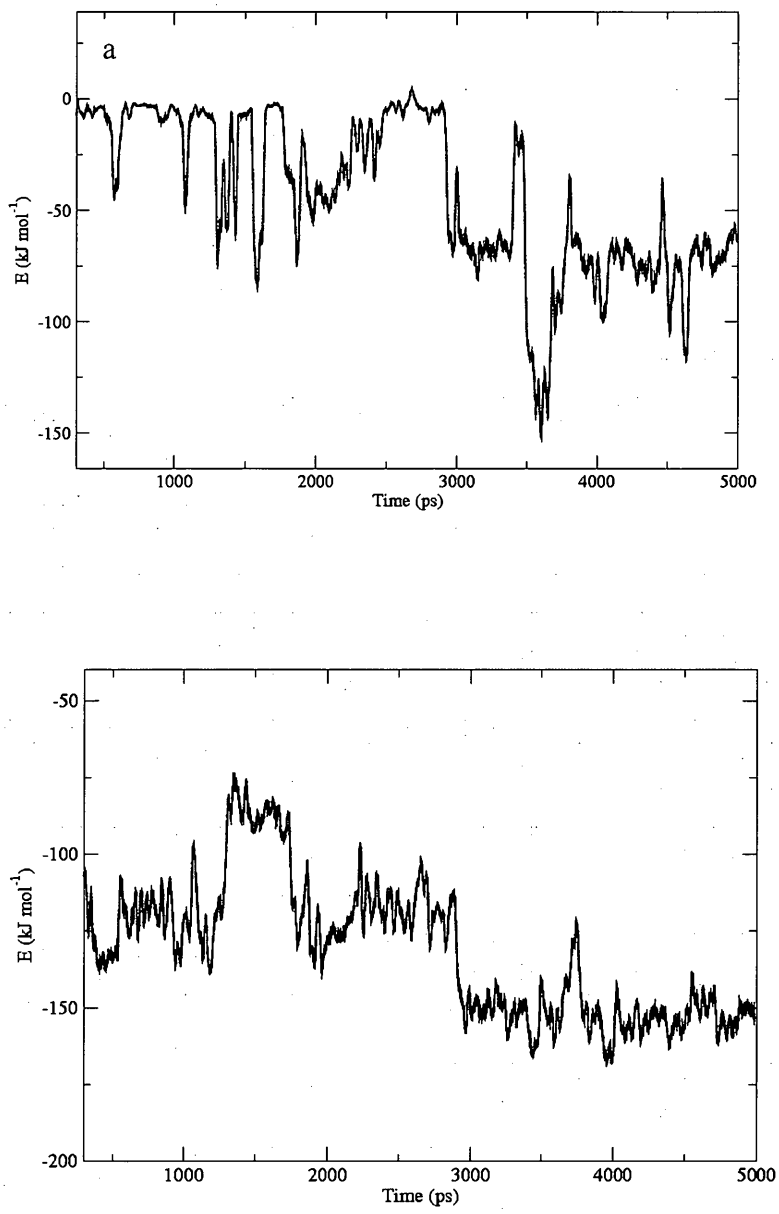


Fig. 4.5 Electrostatic (a) and dispersive (b) potential energy contributions to the overall protein-ligand interaction potential.

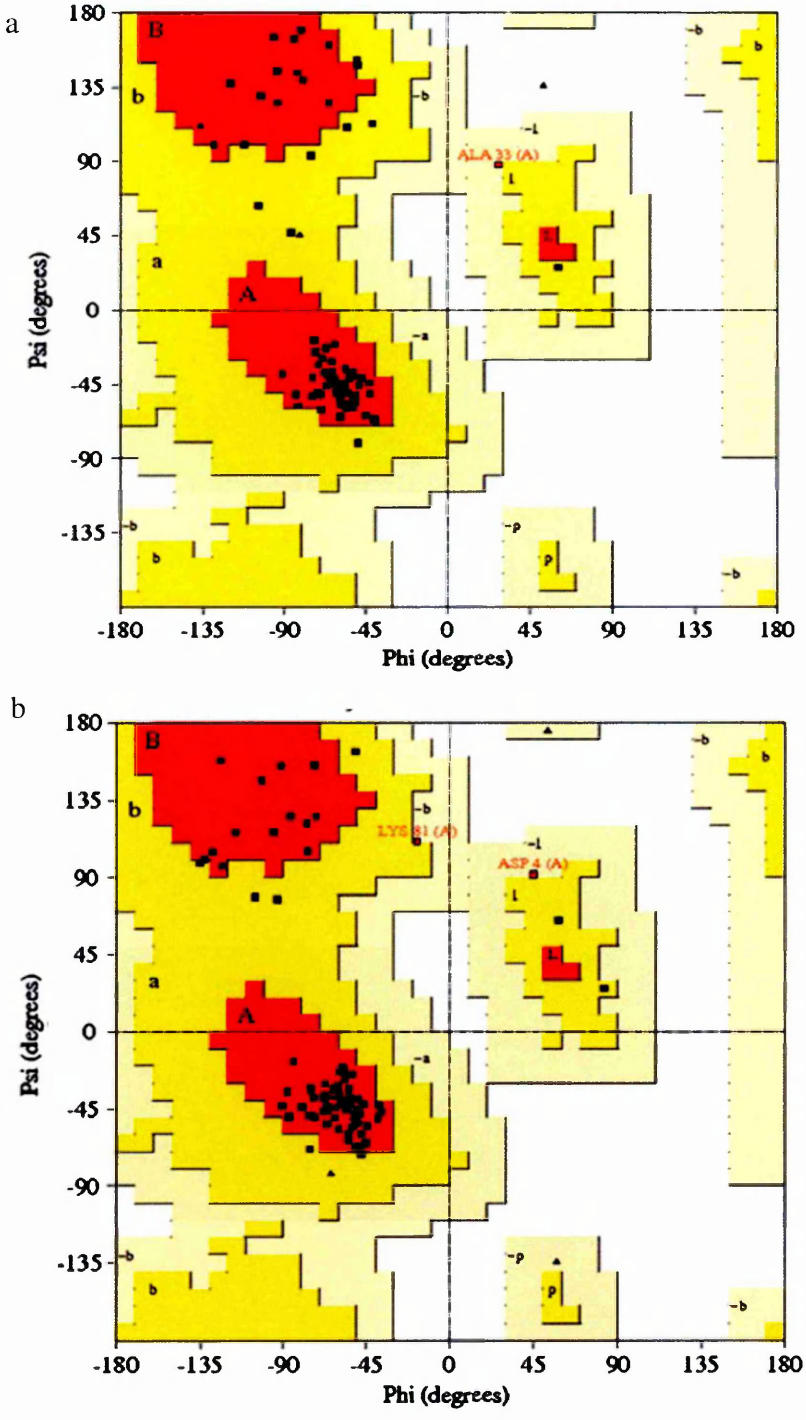


Fig. 4.6 Ramachandran plots of the starting (a) and final (b) protein configurations obtained from the 5 ns MD simulations. The states reported correspond to the ones shown in Fig. Fig.4.3a and Fig.4.3b respectively.

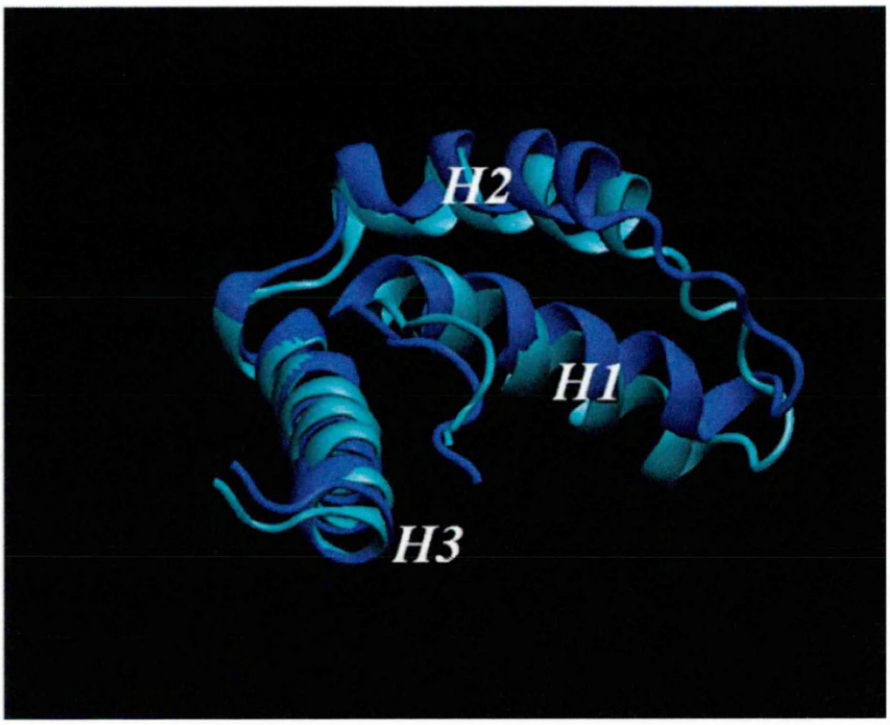


Fig. 4.7 Superimposition of protein structures at the end (blue) and at the beginning of the 5 ns MD simulation.

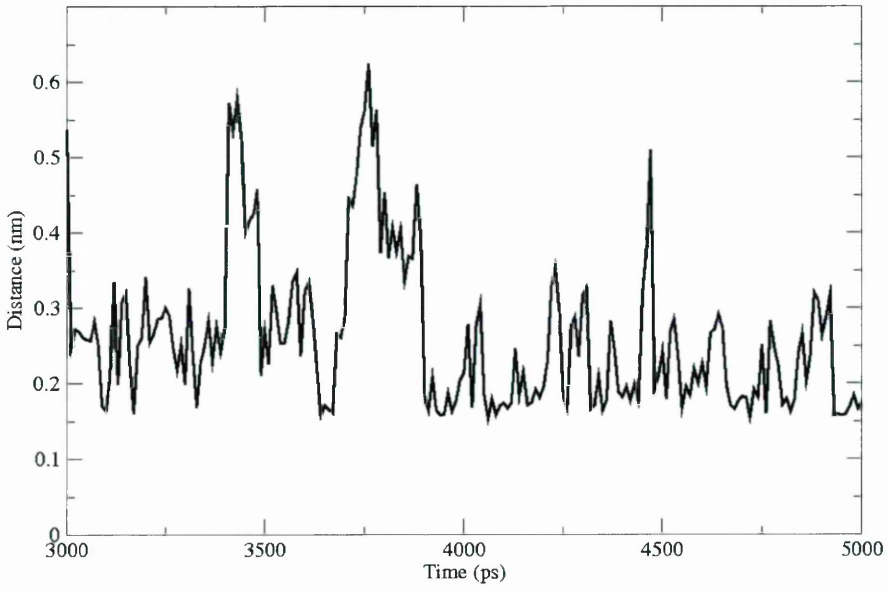


Fig. 4.8 Length of hydrogen bond between Hβ of Ser38 and O34 of CBNX monitored during the 5 ns MD simulation.

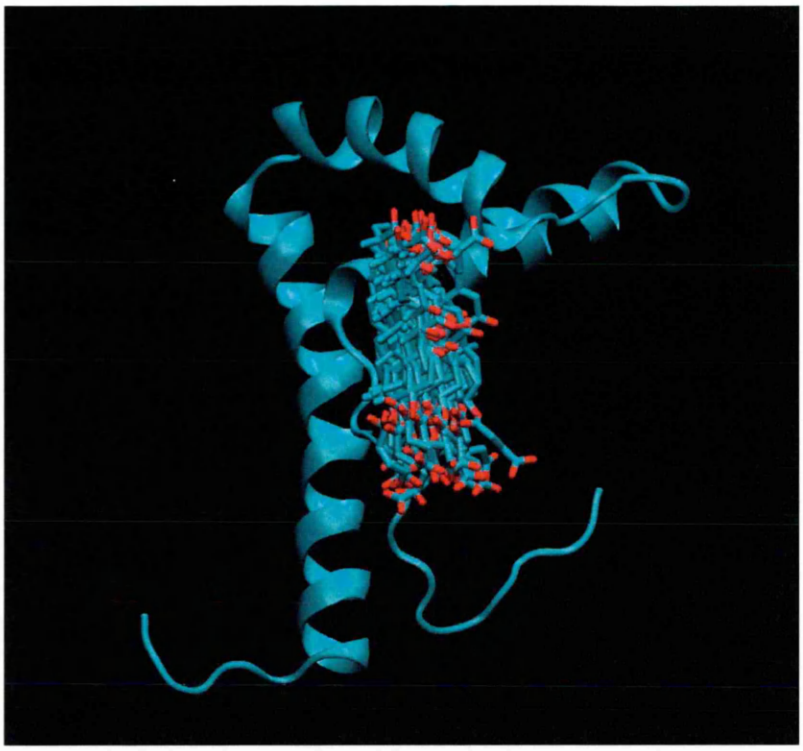


Fig. 4.9 22 Cluster structures for bound CBNX extracted from 3000 frames of NVT REMD simulation. Clusters have been obtained using a 0.1 nm cutoff and applying the linkage algorithm.

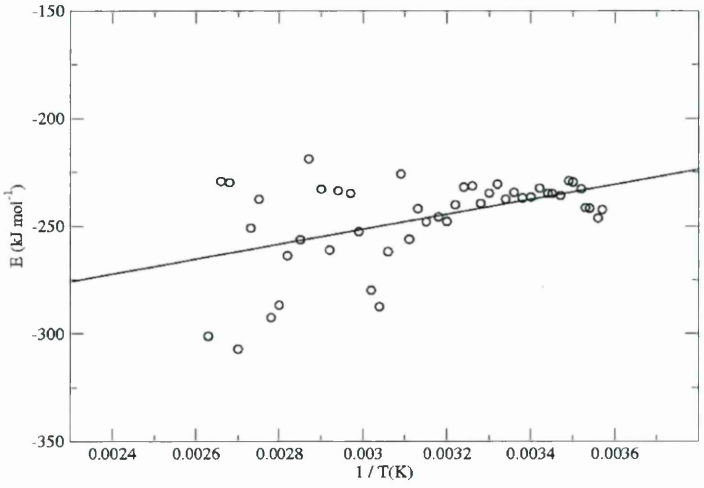


Fig. 4.10 Protein-ligand potential interaction energy expressed in function of 1/T for each replica in NPT ensemble. The energy has been extracted averaging the values sampled in the last 500ps of simulation for each replica.

Fig. 4.11 Two structures of Box A – CBNX complex obtained sampling the 5ns unrefined GROMOS FF based simulation. The one reported in (a) differs from the other in (b) due to the formation of three stable hydrogen bonds. The distances between donor (Hb Ser38) and acceptor (O11 CBNX) are shown for the two cases.In (b) the sidechain is also indicated.

at high temperature the energy distribution becomes scattered but never exhibiting a specific trend. This observation is fully consistent with the negligible contribution of directional H-bonds. In summary, the analysis of the QRFF–MD trajectories leads to the conclusion that several hydrophobic contacts drive the ligand–protein interaction and that the relevant characteristics of the ligand are its shape and its global charge distribution. Therefore these findings confirm that a quantum refinement of the FF was necessary in order to improve the parametrization of long-range electrostatic interactions, also preserving consistency with the van der Waals interaction parameters that are relevant for non directional interactions.

In order to provide a reference for the improvement gained by the usage of the quantum refined force field, also a 5 ns MD and a 60 ns REMD simulations with standard GROMOS force field were performed. Within a 5 ns MD run the ligand fluctuated essentially between two states (reported in Fig. 4.11) of almost comparable

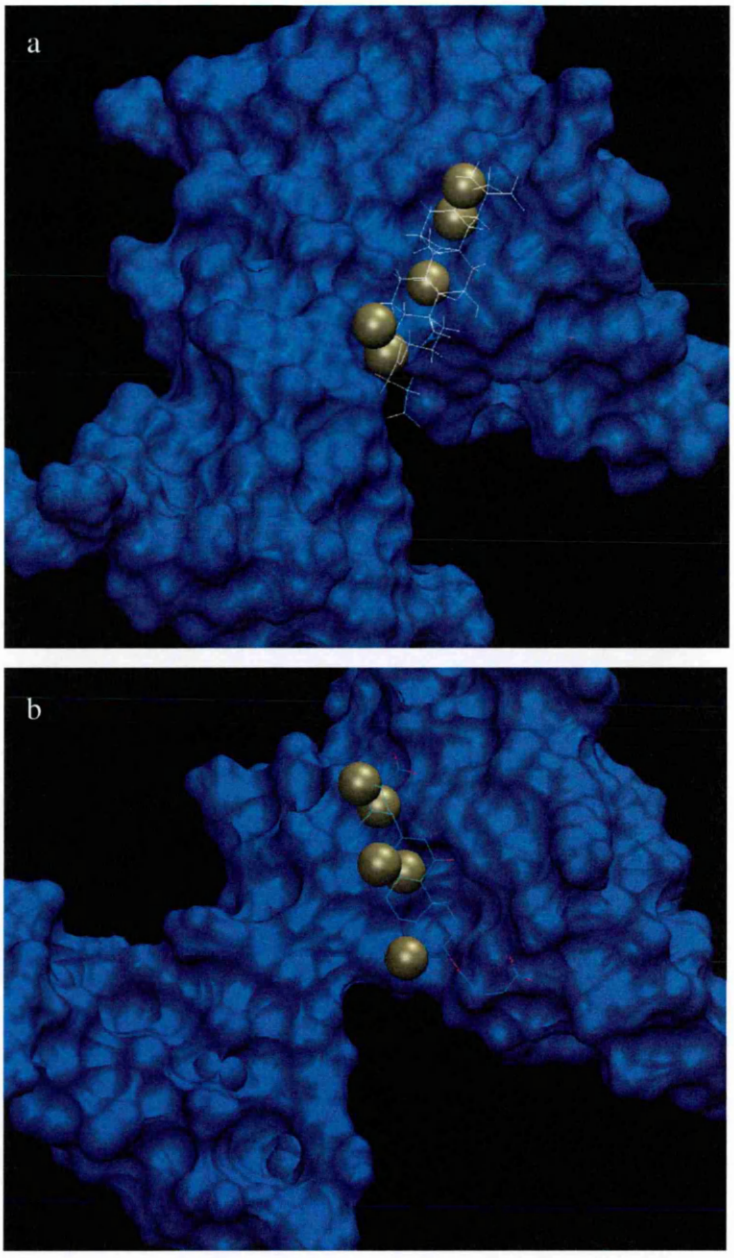


Fig. 4.12 Structures of the Box A–CBNX complex obtained by QRFF (a) and GROMOS (b) based simulations. Protein has been rendered by surface representation, ligands are represented in lines and hydrogen atoms that resulted to be in contact from STD experiments have been reported using brown spheres.

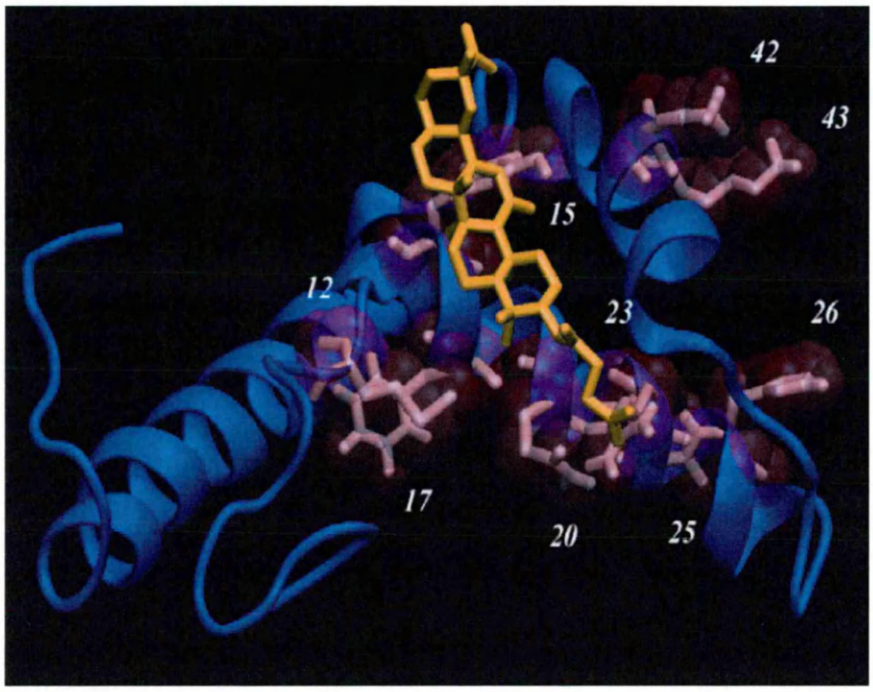


Fig. 4.13 Complex final configuration obtained by REMD simulations adopting unrefined GROMOS FF.

probability characterized by contacts with Tyr15, Ala16 and F37 and by slightly different orientations. In the configuration reported in Fig. 4.11a (59% probability) the C11–O11 carbonyl is directed towards the Helix 2 of the protein and three H-bonds were formed: a short one, O11–Ser41 (1.6 ± 0.1 Å), and two longer ones, O32–Ser38 (2.5 ± 0.9 Å) and O34–Phe37 (2.7 ± 0.6 Å), all with a lifetime on the order of 100 ps. In the other state (41% probability, see Fig. 4.11b), H-bonds were absent. Subsequent REMD search (60 ns) led essentially to stabilize the one reported in Fig. 4.11a: the O11–Ser41 bond was 1.7 ± 0.1 Å long with a lifetime up to 2.5 ns, whereas O32–Ser38 (2.1 ± 0.6 Å) and O34–Phe37 (2.2 ± 0.4 Å) bonds had lifetimes up to 0.3 ns. In this sense, unrefined FF overestimates the directional forces and hence pushes the ligand towards the Helix 2 and does not allow ligand reorientation.

The protein interacting surface and the overall position of the triterpenic ligand scaffold obtained with the two methods have been compared with experimental data (CSD and STD). The structure of the complex obtained from QRFF based simulations

displays excellent agreement with CSD (Fig. 4.3b) and STD (Fig. 4.12) data, whereas the GROMOS based complex structure (*i.e.*, the one obtained by REMD simulations) is not fully consistent with these data (Fig. 4.13): in particular, Met12 and Phe17 CSD remains partially unexplained and also some of the methyls identified by STD (12, 19, 26). In order to further investigate the reliability of the obtained model, NOE peak intensities have been compared to proton-proton distances averaged on equilibrium configurations of the complex sampled from QRFF and GROMOS based simulations. Table 4.2 reports the results of distance calculations for the contacts (Neuhaus et al., 2000) with the two aromatic rings of the Tyr15 and Phe37 side-chains computed using an r^{-6} averaging (Zagrovic et al., 2006) and referring to averages over 8 replicas centered at 310 K for REMD simulations and to the last 2 ns for the 300 K MD simulations. Although the comparison can only be qualitative, global consistency between the experimental observations and the values derived from the QRFF trajectories can be seen, at least for strong and intermediate NOEs. Discrepancies are more significant when the trajectories are taken from the MD runs using the unrefined scheme: especially the contacts with Phe37, which appear to be very relevant, would be lost without quantum chemical refinement.

4.2.3 Experimental measurements on Box A–GL and Box B–GL complexes

The same kind of X-filtered experiments previously described have been performed on Box A–GL and Box B–GL complexes in order to establish the protein-ligand contacts. Also in the present case up to 15 intermolecular contacts between HMGB1 boxes and GL have been found: the residues involved in these contacts were the aromatic residues located at the junction of the two arms of the L-shaped fold of Box A (Tyr15 and Phe37) and of Box B (Phe101 and Phe102). Table 4.2 reports the contacts between the proteins and the ligand and how they have been used as ambiguous and unambiguous restraints.

In order to take into account for local protein distortion induced by the presence of the binding ^{1}H-^{15}N RDCs have been measured using cylinder shaped 6% acrylamide gels (see Materials and Methods). The values experimentally obtained for each backbone amide bond for both free and bound proteins are reported for Box A and for

Box B in Fig. 4.14 and in Fig. 4.15 respectively.

4.2.4 Data driven docking calculations performed on Box A–GL and Box B–GLr complexes

Despite the chemical similarity between GL and CBNX, the application of QRFF methodology presented led us to unreliable results and also to several calculation failures probably due to a partial inconsistency of the novel force field. This can be largely explained by the very well documented difficulty in properly describing sugar rings in the most chemically reliable way adopting classical and semiclassical schemes (Woods et al., 1995; Guvench et al., 2008). The refinement of an existing force field was behind the purpose of the present work.

We therefore adopted the HADDOCK strategy (Dominguez et al., 2003), a recent but well-established procedure which uses experimental chemical-shift data to drive molecular docking, both of proteins and of small molecules (Schieborr et al., 2005). In order to obtain an atomic level model of Box A–GL and Box B–GL complexes all the experimental data obtained for these systems have been used, i.e. chemical shift differences between free and bound proteins (see Chapter 1), intermolecular NOEs, protein RDCs and relaxation data (see Chapter 1).

The usage of RDCs in HADDOCK allowed in both cases the generation of an ensamble of protein structures within the complexes whose backbones are extremely well superimposed, leading to an average Cα - RSMD of 0.41 Å for Box A (min 0.19 Å; max 0.73 Å) and of 0.81 Å for Box B (min 0.28 Å; max 0.63 Å). Q factors computed for Box A and Box B, excluding the unstructured regions, are 0.34 and 0.41 respectively.

HADDOCK docking procedures generated an ensemble of 2000 structures, from which only the best 200 (*i.e.*, according to a minimum energy criterion) were extracted and refined in an explicit water model. These last 200 were clustered according to geometrical superimposition of ligand (see Materials and Methods) with a cut off of 3 Å and a minimum number of 4 structures for bundle (in order to ensure a statistical significance). Clusters populations are reported in Fig. 4.16 and Fig. 4.17. In both cases

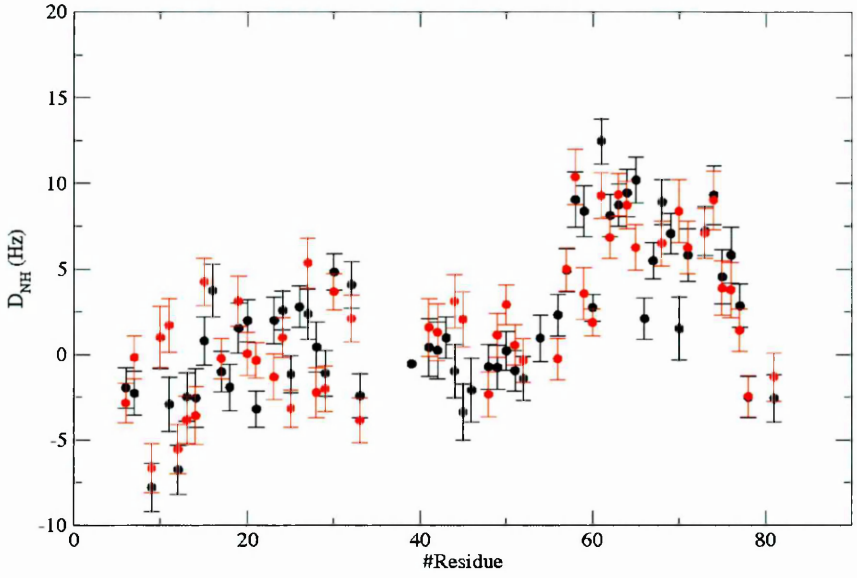


Fig. 4.14 Residual dipolar copling values for amide bond for free (black) and bound (red) Box A.

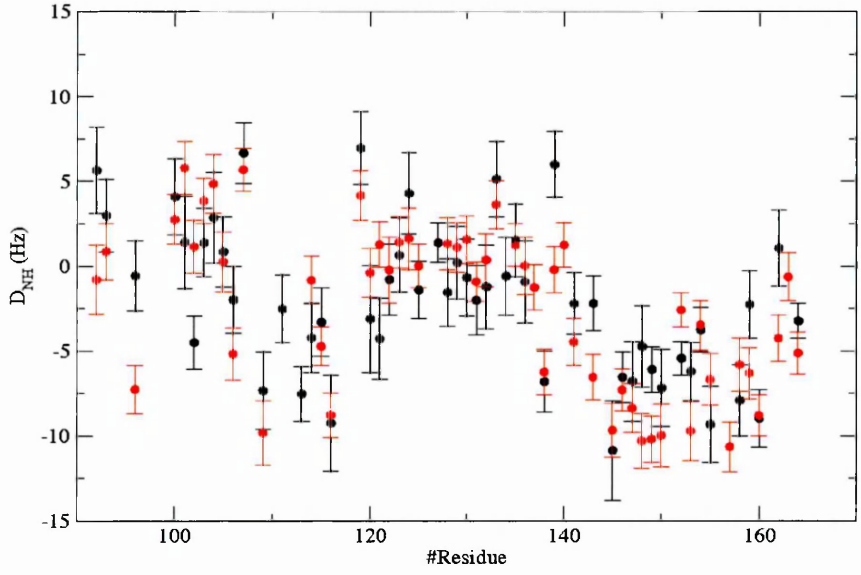


Fig. 4.15 Residual dipolar coupling values for amide bond for free (black) and bound (red) Box B.

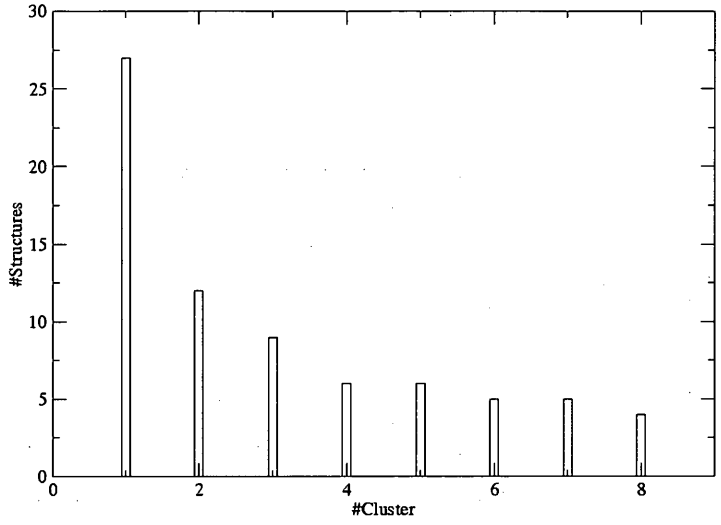


Fig. 4.16 HADDOCK clusters population for Box A–GL complex.

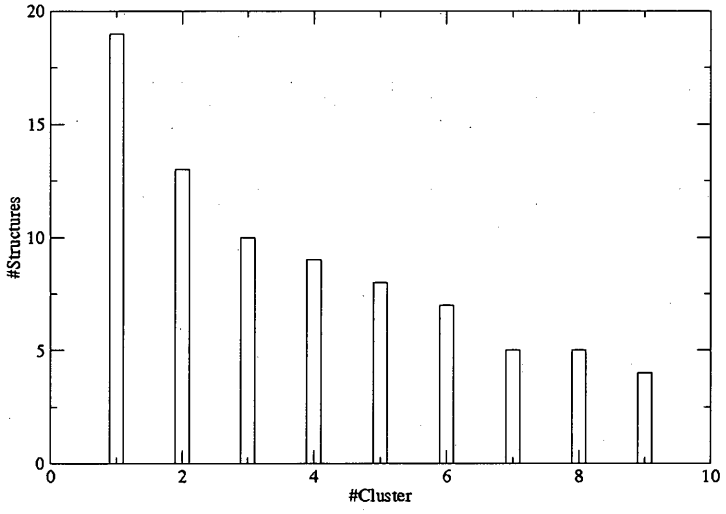


Fig. 4.17 HADDOCK clusters population for Box B–GL complex.

a similar number of clusters have been obtained (8 for Box A, 9 for Box B). The overall number of structures that were successfully clustered is similar (70 for Box A, 72 for

Box B) and the population distributions only slightly differ.

Structures and clusters have been ranked adopting both an energetic and statistical criterion. Clusters have been considered representative for the "true" structure when displayed the lowest interaction energy between the protein and the ligand, *i.e.* they had the lowest non bonded interaction energy. Moreover, when the energetic criterion was too stringent (*i.e.*, the clusters energies were too close to each other) to sharply decide between two ensembles of structures, due to the data driven nature of the present methodology, the most populated ensemble was considered to be the most reliable. Fig. 4.18, 4.19 and 4.20 report respectively the non bonded interactions, the Van der Waals and the Coulombic interactions for Box A–GL and for Box B–GL complexes. According to the choice criterion exposed above, cluster 1 has been considered representative of the binding mode of Box A due to the lowest non bonded energy term and the highest population, whereas cluster 7 has been chosen for Box B–GL complex despite the low population but due to the most favorable non bonded energetic term among the other clusters. Moreover, the two clusters considered displayed no significative deviation from unambiguous and ambiguous restraints, from experimental R_2/R_1 ratios (see Materials and Methods) and from experimental RDCs: the statistics about violations are reported in Table 4.3. The four best structures (*i.e.*, according to the energetics of the complex) for each cluster are reported in Fig. 4.21: in both cases the region of interaction is well defined and the bad superimposition of sugar moiety of GL is in good agreement with experimental data that indicate that only the triterpenic moiety is relevant for the binding.

Data driven docking calculations led to a model for Box A–GL complex comparable to the one identified by means of QRFF methodology and reported in Fig. 4.22: the triterpenic scaffold of GL accommodates in the junction of the two arms in a very similar way to the binding mode of CBNX to Box A and establishing favorable Van der Waals interactions with Ala 16, Tyr15, Val19 and Phe37 and sugars pointing away from the Helices 1 and 2. Moreover, the model predicts the formation of a strong hydrogen bond (2 Å long) between the carboxylate of GL and the hydroxyl group of Ser38, that seems to be the only stable and directional interaction of the complex.

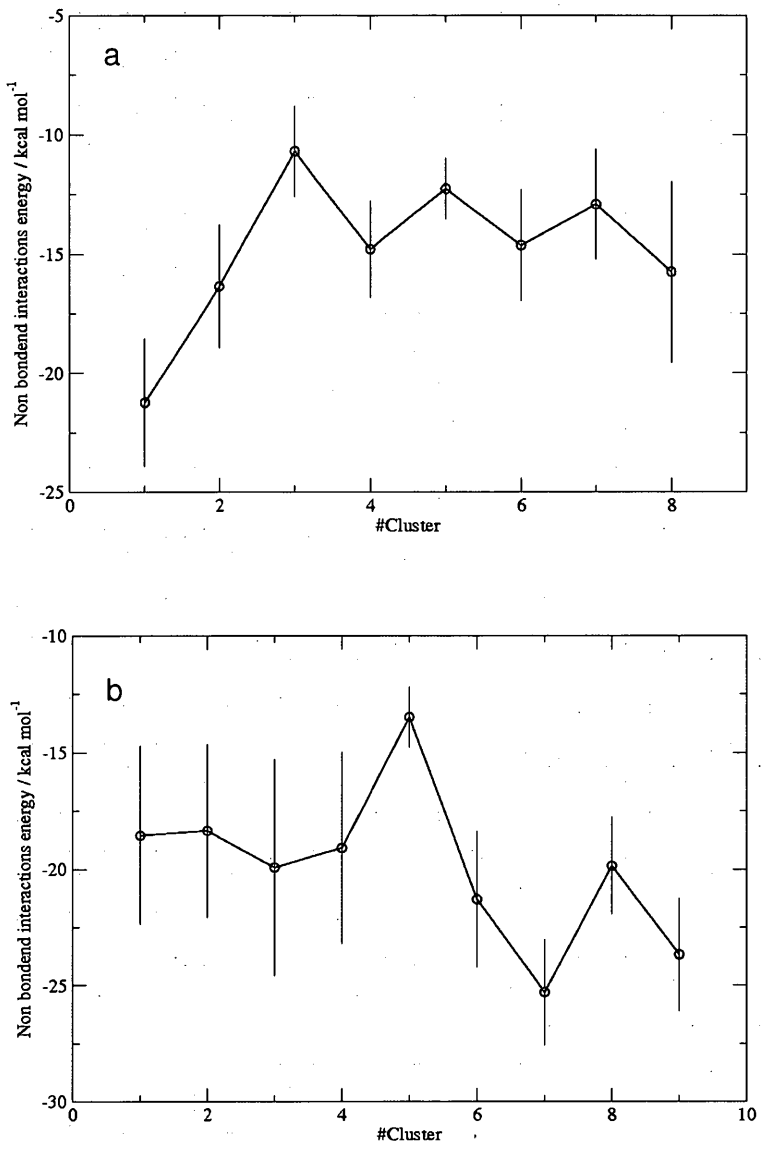


Fig. 4.18 Non bonded interactions energy of Box A–GL (a) and Box B-GL (b) complexes as a function of the cluster number.

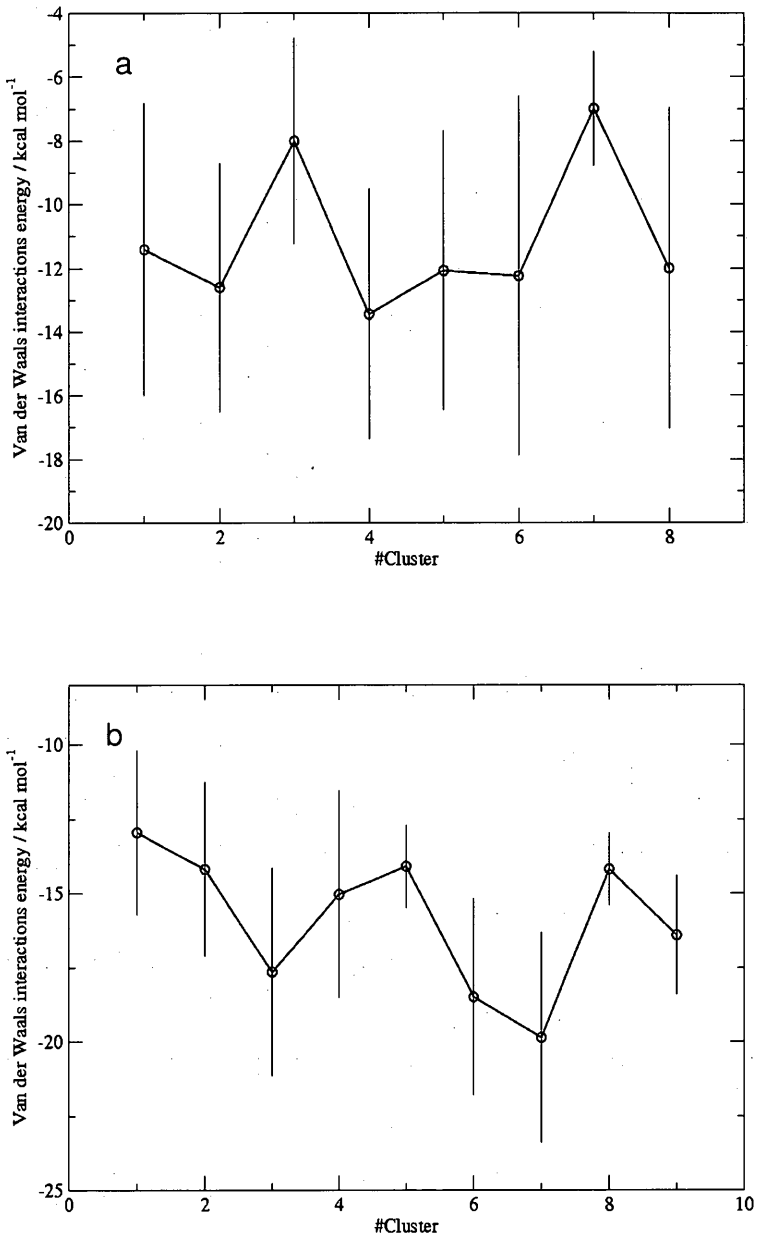


Fig. 4.19 Van der Waals interactions energy of Box A–GL (a) and Box B-GL (b) complexes as a function of the cluster number.

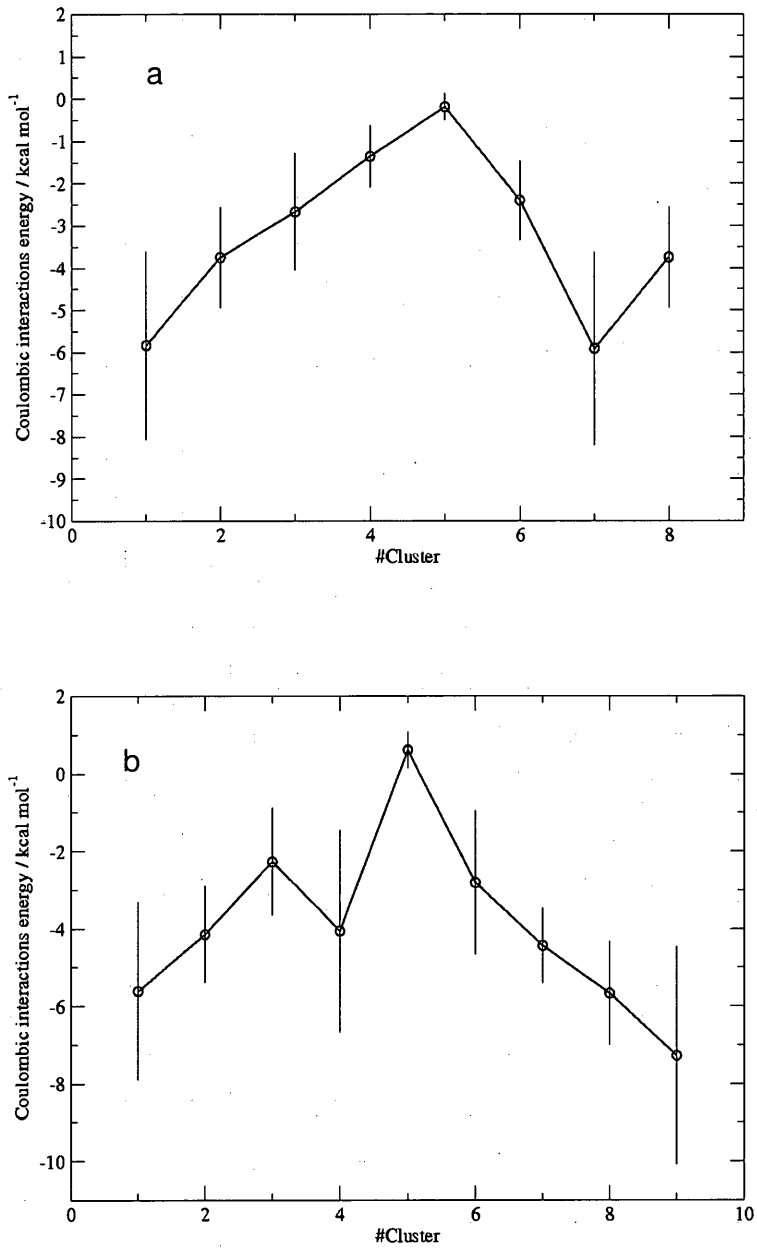


Fig. 4.20 Coulombic interactions energy of Box A–GL (a) and Box B-GL (b) complexes as a function of the cluster number.

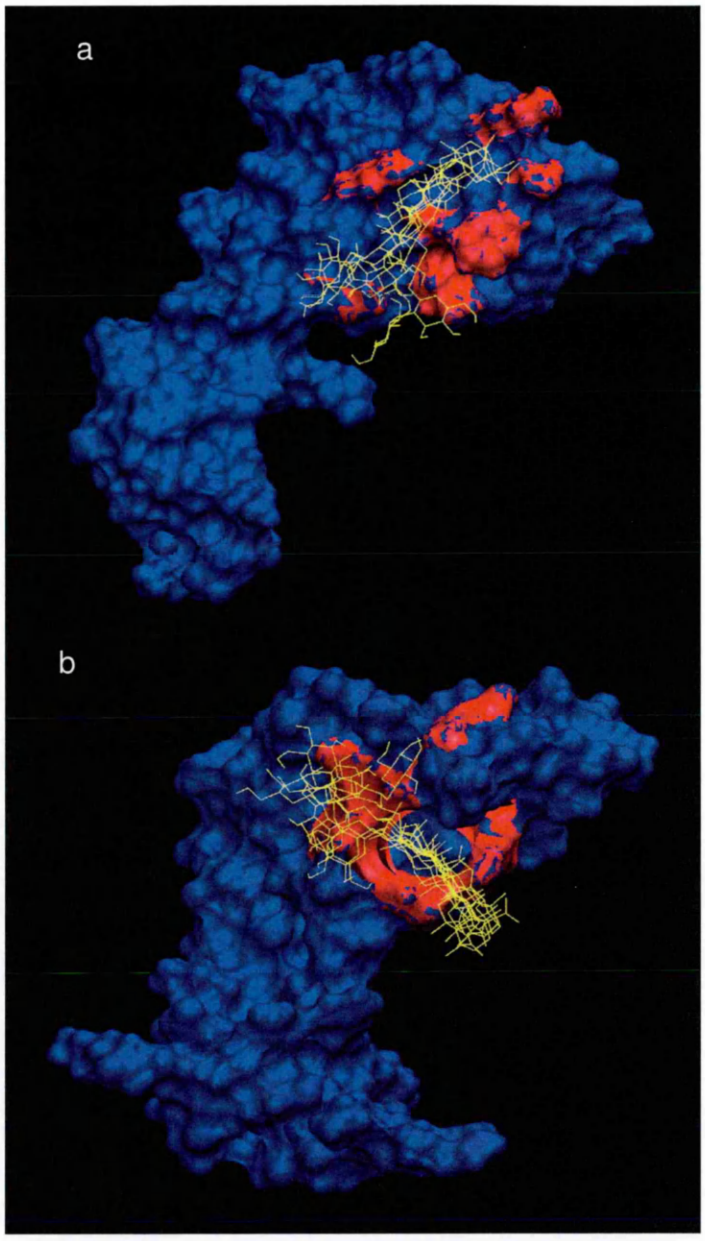


Fig. 4.21 Best 4 structures (i.e., ranked according to the non bonded energy term) of the clusters with lowest non bonded interactions energy for Box A (a) and Box B (b) complexes with GL.

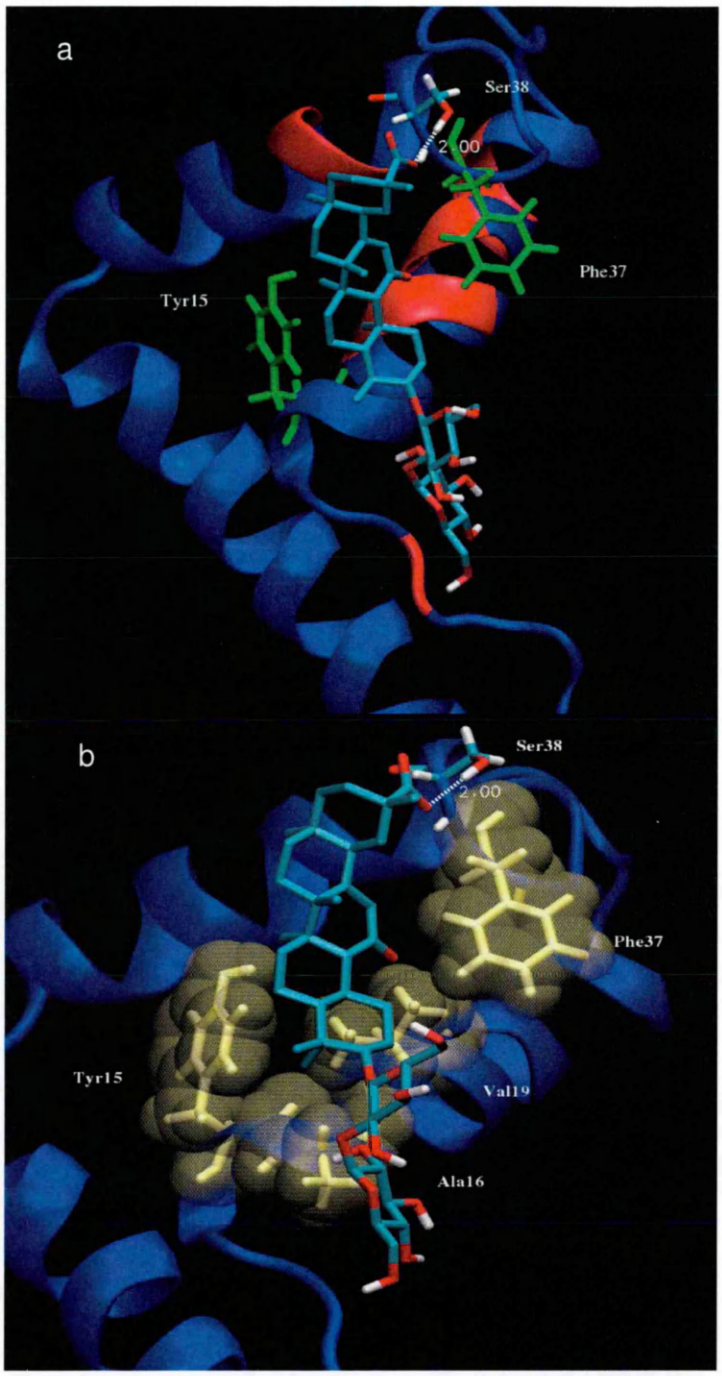


Fig. 4.22 Binding site of GL for the Box A–GL complex. (a) residues with the greatest chemical shift displacement upon binding are reported in red in ribbons representation; the residues with whom NOE contacts have been detected (Tyr15 and Phe37) are reported in green; Ser38 and the bond formed by its sidechain and carboxylic group of GL are reported. (b) Ser38 and the bond formed by its sidechain and GL are reported for sake of clarity.

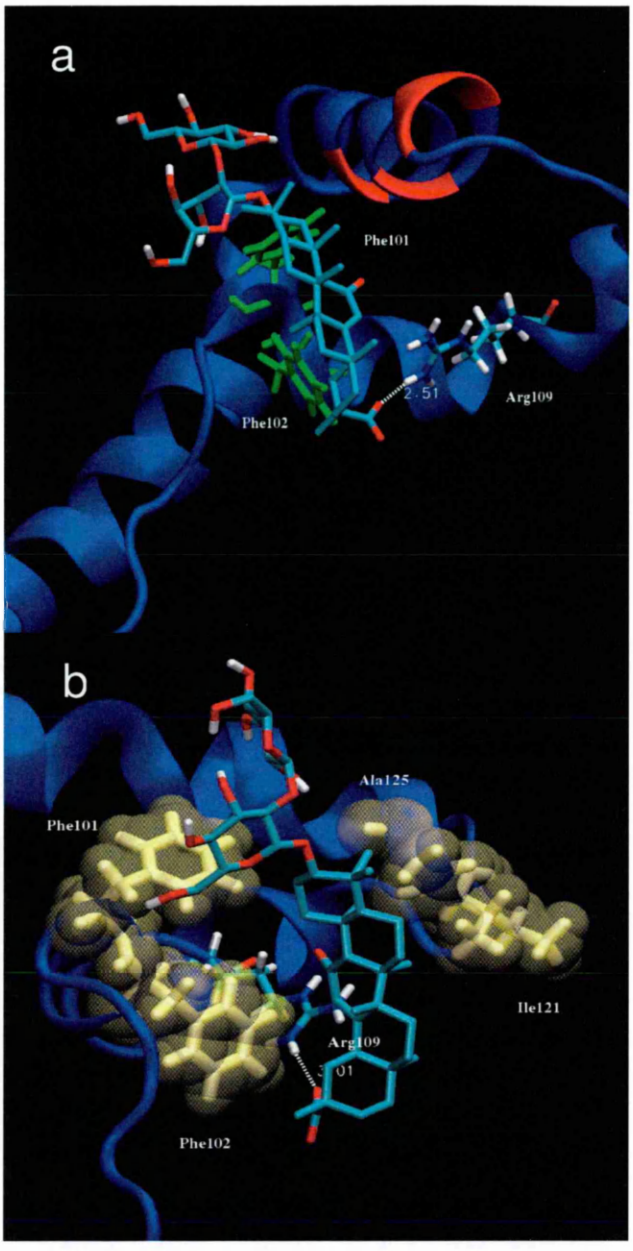


Fig. 4.23 Binding site of GL for the Box B–GL complex. (a) residues with the greatest chemical shift displacement upon binding are reported in red in ribbons representation; the residues with whom NOE contacts have been detected (Phe101 and Phe102) are reported in green ; Arg109 and the bond formed by its sidechain and carboxylic group of GL are reported. (b) Residues that form the hydrophobic stretch that accommodates the ligand are evidenced in yellow through a Van der Waals representation; Arg109 and the bond formed by its sidechain and GL are reported for sake of clarity.

The HADDOCK calculations suggest similar interactions between Box B and GL (Fig. 4.23): the ligand accommodates on the shallow concave surface formed by the two arms of the HMG box, creating favorable hydrophobic interactions between its triterpenic scaffold and the side chains of Phe101,Phe102, Ile121, and Ala125. Similarly to Box A, the model predicts a single directional interaction (an intermediate hydrogen bond, 2.5 Å long) between the carboxylic group of the ligand and the guanidinium moiety of Arg109. Also for Box B the glycosidic moiety of GL points away from Helices and from the binding site.

5. Discussion

5.1 Summary

In the last years the problem of finding novel drugs against inflammation has become central, due to the detrimental effects of known antiinflammatory drugs, whether they are steroidal (SAIDs) or nonsteroidal (NSAIDs) molecules (Rhen et al., 2005). One solution to this scientific and clinical problem is the rediscovery of old drugs that originally were not used against inflammation, like thalidomide, whose side effects are known and well documented. In this perspective, some natural drugs are under investigation, especially molecules whose interactions with the NF-kappaB pathway are documented (like curcumin or epigallocatechin 3 gallate) (Maroon et al., 2006). However, in the last years the main focus of a large part of the scientific community is the identification of pathways and proteins involved in inflammation processes that are new with respect to the "traditional" inflammation mechanisms (Chong et al., 2007). A few years ago it was demonstrated that High Mobility Group Box 1 protein (HMGB1) acts both as an architectural chromatin protein and as a cytokine when released in the extracellular space both actively or during cell necrosis (Muller et al., 2001; Yang et al., 2005), mediating proinflammatory processes like cell migration and proliferation, hence becoming a promising target for antiinflammatory drugs. Recently, it has been demonstrated that glycyrrhizin (GL), a naturally occurring triterpenic drug widely used in clinical trials against SARS and Hepatitis C virus (Coon et al., 2004; Fuji et al., 2004), is able to inhibit HMGB1 induced cell migration and proliferation (see Preliminary results) and to reduce the effect of Hepatitis C in animal models in which it was demonstrated that HMGB1 plays a pivotal role (Sitia et al., 2007). In this sense, the analysis of the physical-chemical determinants of protein-ligand binding is a crucial step for obtaining a full comprehension of the molecular mechanism that regulates the influence of GL on the HMGB1 activity. Moreover, the atomic level description of this interaction is an essential step for any further HMGB1-GL interaction based drug development process.

In the present work we addressed the question whether GL interacts directly with HMGB1 and which role single HMGB1 Boxes have in the ligand binding: moreover, adopting different computational techniques we have described the binding mode of GL and of its analogue carbenoxolone (CBNX) to HMGB1 Boxes at atomic level, partially elucidating the atomic level determinants of anti-inflammatory properties of GL and providing a first step towards the design of new derivatives with improved HMGB1-binding properties.

5.2 General features of the HMGB1-GL/HMGB1-CBNX binding

Through fluorescence and NMR spectroscopy we demonstrated that GL binds to both Boxes with a similar K_d (~150 μm). The Boxes in the full length protein are able to bind ligands independently (with a 1:1 stoichiometry) and in a non cooperative way. Fluorescence and NMR based titration are techniques that work at different concentrations of protein (respectively in the micromolar and in the millimolar or high micromolar range), hence the agreement between the K_d values indicate that the drug-protein binding process is not concentration dependent.

As indicated by circular dichroism spectra, the binding of GL to HMGB1 Boxes does not alter the secondary structure of Box A nor of Box B. Moreover, the protein ligand binding is driven by the interaction between the polipeptide chain and the triterpenic scaffold of GL, as suggested by experiments performed with the GL analogue CBNX, that lacks the glycosidic moiety. This interaction pattern is reminiscent of the one described for DNA polymerase β: in this case as well, carbenoxolone binds on the surface of the protein, optimizing the interactions with the hydrophobic cleft formed by two helices (Hu et al., 2004).

The diffusion mode of the complexes remains unaltered upon binding as demonstrated by the very small variations of τ_c between the free and the bound forms of both Box A and Box B. These data also indicate that the complex formation does not modify the monomeric state of the target protein.

5.3 Indentification of HMGB1 binding site for GL/CBNX

By chemical shift perturbation experiments we identified the GL binding site which is localized at the junction of the two arms of the classical L-shape fold of both HMG Boxes, i.e. only Helices 1 and 2 are directly involved in the binding, whereas Helix 3 is untouched by ligand binding. The same region displays a relevant decrease of ^{15}N T_2 values upon binding, confirming the chemical shift perturbation data and suggesting that the binding site of GL becomes more rigid upon binding on a timescale that goes from higher nanosecond (comparable to the correlation time) to millisecond timescale (comparable to intrinsic spin relaxation timescale). Notably, the binding sites for glycyrrhizin on the HMG Boxes partially overlap with the DNA binding sites, shielding residues such as Arg23, which is well known to be crucial for DNA binding (Ohndorf et al., 1999). The modest effect of glycyrrhizin on the intranuclear function of HMGB1 is in agreement with the absence of cytotoxicity even at high glycyrrhizin concentrations and with the good pharmacological tolerability of glycyrrhizin in rodents and humans (Ploeger et al., 2001). This is possibly due to the vast difference in the affinity of glycyrrhizin and distorted DNA for HMGB1 (on the order of 100 μM and 10 nM, respectively) (Bianchi et al., 1992). Interestingly, glycyrrhizin does not cause the release of HMGB1 from apoptotic chromatin, implying that it will not produce paradoxical proinflammatory responses to apoptotic cells.

5.4 The influence of GL binding on the HMGB1 backbone mobility

The "model free" analysis of relaxation data allowed the determination of K_{ex} values for some residues and the consequent evaluation of exchange contributions for each residues where present, i.e. the exchange of amide protons between environments due to conformational local transitions. Upon binding, the backbone dynamics of residues Ala19, Arg23, Lys27, Asn36, Ser38 and Arg47 located on Box A display exchange contributions (Fig. 5.1a), whereas residues Glu73, Met74 and Thr76 upon binding loose their exchange contributions to overall dynamics on a slow (ms) timescale. The residues located on Helix 1 and 2 can be seen as probes of binding kinetics, i.e. amide backbone

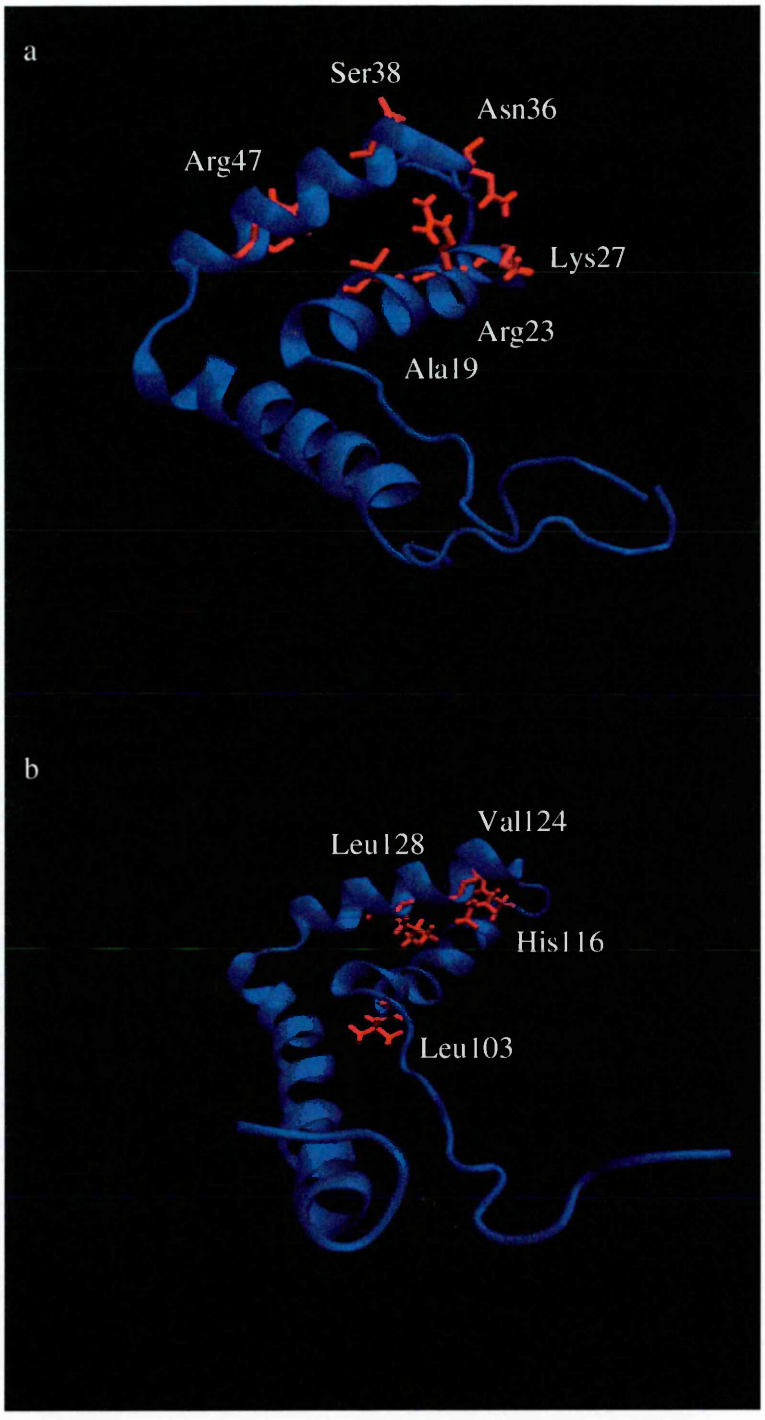


Fig. 5.1 Residues of Box A (a) and Box B (b) that upon binding display a K_{ex} term in the Lipari-Szabo analysis.

protons exchange between bound and free states at a frequency (see Fig. 3.12) of 5-10 Hz. The loss of exchange contributions for residues located on the C terminal part of Helix 3, even though they are not involved in direct contact with the ligand, indicates that the binding of GL influences all the protein dynamics and perturbs the conformational sampling of residues that are not located in the binding site. The same behavior has recently been observed in Heat Shock Protein 90 (Colombo et al., 2008) where ligand binding is reflected by enhancement or freezing of protein motions also distant from the binding site. In the cited reference, the authors advanced the hypothesis that a protein can exhibit differences in the local dynamics as a function of a binding partner, and that molecular recognition is to some extent compatible with well determined and possible motions. In this sense, our findings suggest that GL binding freezes slow motions of the C terminal part of Box A, that becomes more tethered and hence displays a different accessibility (with respect to free protein) towards other proteins or ligands: in principle this could allow interactions that are mediated by GL binding even without a direct contact with the small molecule.

In the Box B–GL complex the dynamics perturbation is highly localized on Helix 2 only (K_{ex} is increased upon binding, with respect to protein only, for residues Leu103, His116, Val124 ans Leu128, Fig.5.1b), indicating that this helix is more sensitive to the presence of the ligand. Interestingly, no long range effect is observed like in the case of Box A; residues Ala136, Tyr143 and Lys146 no longer display exchange in the bound state, indicating that the interhelical loop region is stabilized by the binding. Moreover, fast motions in Box B are at least partially blocked by the binding, as indicated by an average increase of S^2 parameter through the whole sequence.

5.5 Atomic level description of the binding

The model structures of the complexes formed by GL and HMGB1 Boxes, obtained adopting a data driven docking procedure as implemented in HADDOCK, were able to provide an atomic level description of the binding incorporating all the experimental data in structure calculations and refinement. HADDOCK is able to generate many structures (ensembles) starting from different sets of experimental data; however, a

complex structure generated adopting this procedure cannot be considered reliable as a standalone structure, but within an ensemble of structures that share one or more common features according to a clustering criterion. In the present study the criterion adopted was the one of the lowest non bonded average interaction energy, followed by cluster population in case of similar cluster energies. For all the GL complexes studied, the cluster with the lowest non bonded energy displays almost the same non bonded average energy (-22 kcal/mol and -25 kcal/mol for Box A–GL and Box B–GL respectively): the most energetically favorable cluster for Box A coincided with the most populated one (cluster 1), whereas Box B–GL most favorable cluster is one of the less populated (cluster 7, see Fig.4.17 and 4.18). However, the first ranked and most populated cluster obtained for Box B–GL complex has a configuration (Fig. 5.2) very similar to the one displayed for the cluster that has the most favorable non bonded interaction energy and that has been chosen as representative of the binding mode of GL to Box B (Fig. 4.21b) despite a reduced population. This information together with the energy profile of clusters energies (Fig. 4.18), indicates that the GL binding region presents a quite flat energy landscape, that allows the accommodation of slightly different configurations of the ligand with a similar energetic content: among them, some configurations are more populated and slightly more favorable than others. In this sense the result of the presented docking calculations can be statistically interpreted as representative of the mobility of the ligand within the binding site. This interpretation recalls the principle of a drug discovery techniques that has recently defined as 4D QSAR (Lukacova et al., 2004), where the binding mode of a ligand takes into account the binding site of a protein, the 3D structure of the ligand and the local mobility of the ligand itself within the binding site: in this way, not only the classical Hansch lock-and-key analogy is used (Hansch et al., 1964), but also local geometry variations are taken into account for exploring local minima.

For both Box A and B the ligand accommodates in the same region of the protein, that corresponds to the HMGB1 DNA binding region (Ohndorf et al. 1999; Read et al. 1993; Weir et al., 1993) adopting a different orientation. Single specific electrostatic interactions between the carboxyl group of the triterpenic scaffold of the

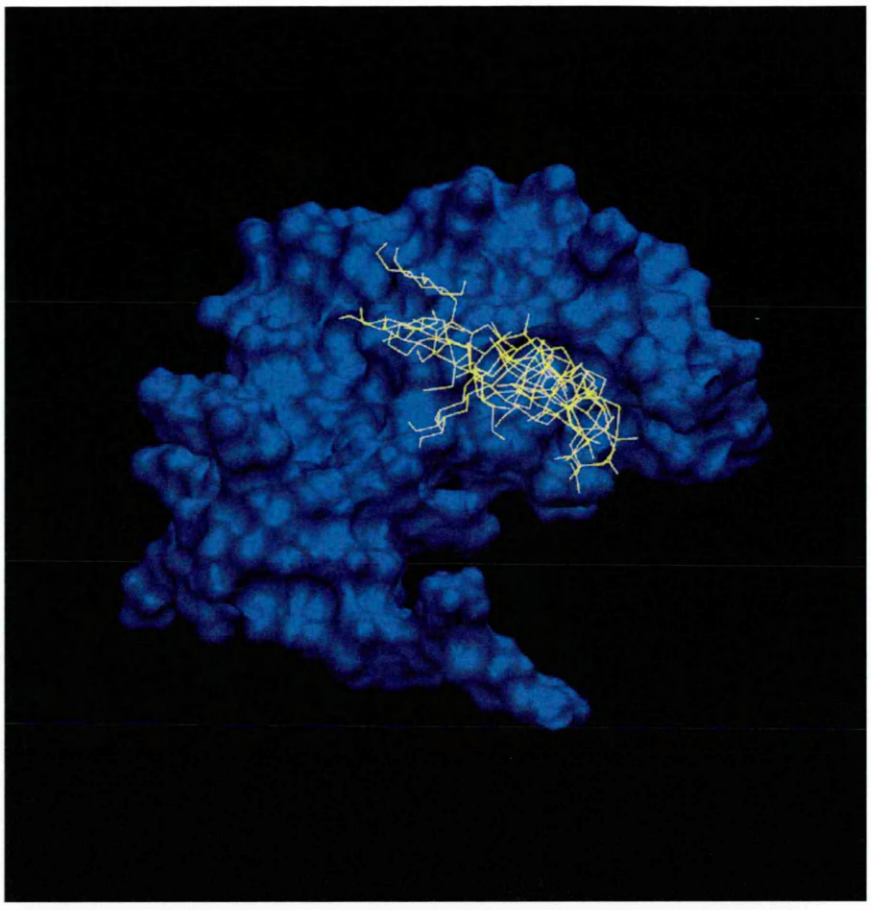


Fig. 5.2 Best 4 solutions (accoding to non bonded energy of the complex) from the cluster 1 obtained adopting the HADDOCK strategy for the Box B-GL complex. Comparison has to be done with Fig.4.21b, that reports the most energetically representative but less populated cluster 7.

ligand (O34) and the sidechain of a single residue can be observed: the interacting residue is located on the Helix 1 (Ser38) and on the Helix 2 (Arg109) for Box A and Box B respectively. The different ligand orientations obtained suggest that, although Box A and B share a similar fold, the region that is able to perform molecular recognition has different chemical features in the two domains, allowing different binding mode of the same ligand. This hypothesis is in agreement with experimental works (based upon electrophoresis shift mobility assays) that demonstrated that Box A and Box B have different affinity towards some DNA forms like supercoiled and 4WJ DNA (Box B can bind with equal extent to the two forms, whereas Box A–4WJ DNA complex is more

stable) (Poheler et al, 1997). Moreover, the chemical difference between the two binding sites could at least partially explain the different chemoattractant activity of Box A and Box B (Degryse et al., 2001) and, in the present work, the different behavior of Box A and Box B when complexed with low solubility ligand like CBNX.

In Chapter 4 a different strategy, with respect to docking, based on *ab initio* refinement of the standard force fields charges has been used to determine the structure of the complex formed by CBNX and Box A. The choice was twofold: i) CBNX lacks glicosidic moieties, that are very difficult to handle with quantum mechanics tools; ii) CBNX induced Box B precipitation even at low molar ratio, impairing the collection of structurally relevant information about this complex. This lead us to adopt for Box A-GL and for Box B-GL complexes the HADDOCK based strategy described before. However, QRFF based structure determination can be viewed as a valid tool for drug discovery (Ferrara et al., 2006) and for almost blind structure determination starting from a few experimental data. QRFF based method moves from first physical principles with the only help of chemical shift mapping results for positioning the ligand in the correct protein site, then allows system evolution towards minimum energy states using classical schemes. HADDOCK instead makes use of all the experimental data for generating the structure, adopting an energy minimization procedure that is very similar to standard NMR based structure determination. There is a good agreement between the QRFF and the HADDOCK determined structure of Box A–ligand complexes (FIG.5.3a) for what concerns the overall ligand orientation, although the two complexes cannot be directly compared. In both cases an interaction between the sidechain of Ser38 and a carboxyl group of GL/CBNX has been predicted (Fig. 5.3b), as expected for similar ligand in the same binding site of a protein, and the orientation of the two ligands is similar with respect to the overall position of the triterpenic moiety of GL and CBNX. These findings suggest that QRFF based methodology is a valid tool for the determination of structures in a drug discovery process and for the screening of many ligands (e.g., a ligands library). Our calculations indicate that this *ab initio* methodology, even starting from few experimental data, can lead to reliable results

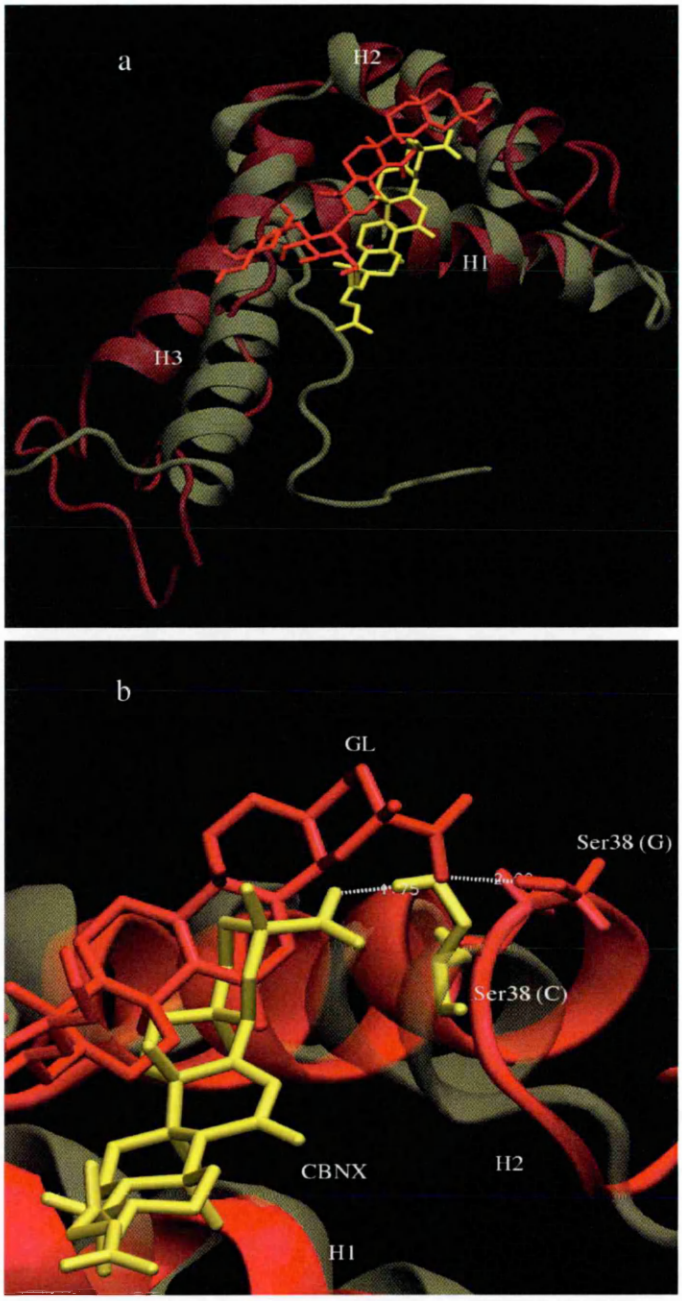


Fig. 5.3 Binding modes of GL (red) and CBNX (yellow) to Box A according to HADDOCK and QRFF based strategies respectively. a. Overview of the binding b. Interaction between the ligand and the sidechain of Ser38.

Despite the goodness of the results obtained and the exiguous experimental data used, the QRFF methodology is computationally expensive and it can only be applied with success and usefully if supported by a massive parallel computing frame. The HADDOCK based strategy, on the other side, has the disadvantage that a quite large amount of experimental data must be provided for getting good convergence in finding the solution to the problem of protein-ligand structure determination bypassing the NOE data collection: however, the very simple implementation of the software and the high number of structures that can be generated in a small amount of computing time make this approach suitable for getting a reliable and fast model of the complex under investigation.

5.6 Biological relevance of HMGB1-GL/CBNX interaction and future perspectives

The chemotactic and mitogenic properties of HMGB1 depend on its interaction with the receptor of advanced glycation end products (RAGE) (Palumbo et al 2004), even if the real nature of this interaction still remains controversial (Bierhaus et al., 2006). It is interesting to note that the RAGE binding surface on HMGB1 is represented by the stretch of basic amino acids between Box B and the acidic tail (Palumbo et al. 2004), and does not correspond to the surfaces engaged by GL. Moreover, the IC50 for the inhibition of HMGB1-induced chemotaxis by GL is 50 μM for fibroblasts (see Introduction, Preliminary Data subsection, Fig.1.12 and 1.13), whereas the dissociation constant for GL and HMGB1 is around 100-150 μM. These results suggest that another surface molecule, possibly belonging to the integrin family as recently suggested (Chavakis et al. 2003), also takes part in the binding of HMGB1 to RAGE, and favors the binding of GL to HMGB1 on the cell surface. Glycyrrhizin is known to interfere with 11b-hydroxysteroid dehydrogenase type 2, leading to an enhanced mineralcorticoid effect of cortisol (Ploeger et al., 2001). However, we tested the effect of cortisol on cell migration induced by HMGB1 and fMLP, and found none (data not shown). This suggests that the effects on 11b-hydroxysteroid dehydrogenase type 2 and cell migration are distinct and separate.

The presented findings indicate that extracellular HMGB1 role in inflammation could in principle reveal a new way to address the problem of finding new and safest strategies for blocking inflammation processes. This new view of inflammation processes is very similar to the research performed on IL33 and its interactions with SR2 receptors (Schmitz et al. 2005), that demonstrated that a new way of rethinking inflammation mechanisms and their inhibition is possible by identifying small inhibitors of these nuclear proteins that are released in the outer cell space and act as cytokines. The advantage of adopting an HMGB1-GL interaction based strategy relies on the fact that here a small molecule that binds directly HMGB1 can be proposed as a possible candidate as an antiinflammatory drug (*i.e.*, a lead compound). Moreover, GL has been extensively studied in the past and used in clinical trials, hence its biological pathways are quite well known and characterized and effects on humans are thoroughly documented.

Despite the known features of the chemical and biophysical determinants of the proinflammatory activity of HMGB1 and the presented results about the role of GL and CBNX in inhibiting its cytokine activity in vitro, some questions are still open. Even if no pathological mutants of HMGB1 are known up to date, the usage of mutants of the protein can be of great interest in elucidating the role of some residues in driving the interactions with ligands and in confirming the effective role of HMGB1-GL binding in modulating proliferation and migration. The first step towards this type of investigation could be the determination of a set of mutants that are both able to retain the correct folding of the HMGB1 Boxes and, at the same time, to alter or cancel the GL binding properties of the protein. This can be done with the help of techniques presented within the text (NMR, MD) or recently introduced *ad hoc* for trying to establish the energetics of proteins using simple empirical models (e.g., FoldX)(Guerois et al., 2002), and also with the help of other biophysical techniques like Differential Scanning Calorimetry (DSC) able to determine the stability of mutants. In principle, this first step would lead to a set of proteins (mutHMGB1) that could be ranked with respect to the ability of binding GL in solution as tested by means, for example, of fluorescence spectroscopy.

The second step would be based on the migration and proliferation assays performed as described in the Introduction starting from the two "extremes" of the aforementioned ranking (*i.e.* the mutant with the lowest K_d and the one with the highest K_d or no binding at all), measuring the migration and proliferation of cells using mutHMGB1 as chemoattractant and as proliferative agents and GL as inhibitor of proliferation and migration. These experiments could address the following questions: a. if the mutations (and the subsequent modification of HMGB1 exposed surface) are able to alter the chemoattractant properties of HMGB1; b. if the residues mutated are really modulating the antiinflammatory properties of HMGB1 with respect to a different GL binding affinity. The first step, incidentally, could also identify regions of the protein surface that could be in contact with the HMGB1 receptor(s): this issue can be also covered performing the proliferation and migration assays with mutants that retain the correct protein folding and do not alter the binding mode of GL, *i.e.* mutants of residues that are distant from the GL binding site.

A possible future development of the presented work on the side of the physical chemical issues covered during the thesis work is the setup of an *in silico* drug discovery project that moves from the results of the presented QRFF based calculations performed on CBNX towards a complete screening of a very small library (no more than 100 compounds) of GL-like compounds in order to find reliable targets with a more suitable K_d (i.e., in the nanomolar or low micromolar range) for pharmaceutical purposes. Of course, such a process could only be realized adopting massive parallel computing technologies and developing large scripts that would allow an extended automation of the whole process, from the creation of QRFF charges to the MD simulation analysis. A complete calculation performed on last generation cores equipped with the fastest intercommunication systems would take* no more than 8-12 hours for each compound for a single HMGB1 Box (2 hours for QRFF charges optimization, 6-10 hours for explicit solvent MD calculations). This procedure is to some extent more

*This estimate has been done hypothesizing the usage of no more than 16 cores and 20 configurations for each compound, performing short (~ 500 ps – 1 ns) MD simulations on last generation machines.

computationally and time demanding than other rougher but faster algorithm (rigid body docking, Perola et al. 2004): however, at this stage of the work no experimental information would be needed, and the resulting ligand-protein complex would be physically reliable. In this way a smaller subset of 4 or 5 candidates with the lowest protein-ligand interaction energy would be available: this would be and an excellent starting point for more classical (3/4/5D QSAR, see Lukacova et al., 2004; Lill, 2007) and faster massive screening of large libraries (1000-2000 compounds) of molecules whose chemical structure could be finely tuned on the results of the previous calculations.

This drug discovery process of course should be coupled with NMR based screening experiments like Water LOGSY (Dalvit et al., 2003) or Saturation Transfer Difference (Mayer et al., 1999) performed on the most promising compounds in order to identify a pool of good targets for biochemical experiments that could elucidate the role of HMGB1 in inflammatory processes and to cross validate *in silico* screening. Experimental efforts also should be dedicated, at the end of both the computational and the spectroscopic drug discovery process, to the determination of the effects of the new compounds identified on both proliferation and migration in order to assess the *in vitro* role of a decreased Kd. This can be easily performed using the same assays previously used to describe the antiinflammatory features of GL on the citokine activity of HMGB1. The new compunds could be, finally, tested by means of proliferation and migration assays as described, both with wtHMGB1 and with the most biologically relevant mutants previously selected on the basis of GL based experiments.

6. References

Aktas O, Prozorovski T, Smorodchenko A, Savaskan NE, Lauster R, Kloetzel PM, Infante-Duarte C, Brocke S, Zipp F. *Green tea epigallocatechin-3-gallate mediates T cellular NF-kappa B inhibition and exerts neuroprotection in autoimmune encephalomyelitis.* J Immunol., 2004, 173(9), 5794-800.

Aue D, Bartholdi F, Ernst RR, *Two dimensional spectroscopy. Application to Nuclear Magnetic Resonance*, Journal of Chemical Physics, 1976, 64, 229-2246

Agresti A, Bianchi ME, *HMGB proteins and gene expression*, Curr Opin Genet Dev. 2003, 13(2), 170-8.

Agresti A, Scaffidi P, Riva A, Caiolfa VR, Bianchi ME, *GR and HMGB1 interact only within chromatin and influence each other's residence time*, Mol.Cell, 2005, 18, 109–121.

Barbato G, Ikura M, Kay LE, Pastor RW, Bax A, *Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible.* Biochemistry, 1992, 31(23), 5269-78.

Barnes PJ, *Theophylline in chronic obstructive pulmonary disease: new horizons,* Proc Am Thorac Soc., 2005, 2(4), 334-9.

Barnes PJ, Ito K, Adcock IM, *Corticosteroid resistance in chronic obstructive pulmonary disease: inactivation of histone deacetylase.* Lancet., 2004, 363(9410), 731-3

Bartels CB, Xia TH, Billeter M, Güntert P, Wüthrich K, *The program XEASY for computer-supported NMR spectral analysis of biological macromolecules,* J. Biomolecular NMR, 1995, 5, 1-10.

Baumann E, Stoya G, Völkner A, Richter W, Lemke C, Linss W, *Hemolysis of human erythrocytes with saponin affects the membrane structure*, Acta Histochem., 2000, 102(1), 21-35.

Bax A, Davis G, *MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy*, J. Magn. Reson., 1985, 65, 355–360.

Berendsen HJC, van der Spoel D, van Drunen R, *GROMACS: A message-passing parallel molecular dynamics implementation*, Comp. Phys. Comm., 1995, 91, 43-56.

Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J, *Interaction models for*

water in relation to protein hydration. In: Pullman, B. and Reidel D, Intermolecular Forces, 1981, Publishing Company Dordrecht, 331–342.

Berendsen HJC, Postma JPM, DiNola A, Haak JR, *Molecular dynamics wit coupling to an external bath*, J. Chem. Phys., 1984, 81, 3684–3690.

Bianchi ME, Falciola L, Ferrari S, Lilley DM, *The DNA binding site of HMG1 protein is composed of two similar segments (HMG boxes), both of which have counterparts in other eukaryotic regulatory proteins,* EMBO J., 1992, 11(3), 1055-63.

Bierhaus A, Stern DM, Nawroth PP, *RAGE in inflammation: a new therapeutic target?*, Curr Opin Investig Drugs, 2006, 7(11), 985-91.

Bonaldi T, Talamo F, Scaffidi P, Ferrera D, Porto A, Bachi A, Rubartelli A, Agresti A, Bianchi ME, *Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion*, EMBO J. 2003, 22(20), 5551-60.

Bonizzi G, Karin M, *The two NF-κB activation pathways and their role in innate and adaptive immunity*, Trends Immunol., 2004, 25, 280–288.

Breeze A, *Isotope-filtered NMR methods for the study of biomolecular structure and interactions* Progress in Nuclear Magnetic Resonance Spectroscopy, 2000, 36(4), 323-372 .

Breuner CW, Orchinik M, *Plasma binding proteins as mediators of corticosteroid action in vertebrates*, J Endocrinol., 2002, 175(1), 99-112.

Broadhurst RW, Hardman CH, Thomas JO, Laue ED. *Backbone dynamics of the A-domain of HMG1 as studied by 15N NMR spectroscopy*, Biochemistry, 1995, 34(51), 16608-17.

Caldwell B, Aldington S, Weatherall M, Shirtcliffe P, Beasley R, *Risk of cardiovascular events and celecoxib: a systematic review and meta-analysis*. J R Soc Med., 2006, 99(3), 132-40.

S. Calogero, F. Grassi, A. Aguzzi, T. Voigtländer, P. Ferrier, S. Ferrari, M.E. Bianchi, *The lack of chromosomal protein HMG1 does not disrupt cell growth, but causes lethal hypoglycaemia in newborn mice. Nat. Genet.,*1999, 22, 276–280.

Chen PC, Wheeler DS, Malhotra V, Odoms K, Denenberg AG, Wong HR. *A green tea-derived polyphenol, epigallocatechin-3-gallate, inhibits IkappaB kinase activation and IL-8 gene expression in respiratory epithelium, Inflammation.* 2002, 26(5), 233-41.

Clemett D, Goa KL, *Celecoxib: a review of its use in osteoarthritis, rheumatoid*

arthritis and acute pain, Drugs, 2000, 59(4), 957-80

Chavakis, T, Bierhaus A, Al-Fakhri N, Schneider D, Witte S, Linn T, Nagashima M, Morser J, Arnold B, Preissner KT, Nawroth PP, *The pattern recognition receptor (RAGE) is a counterreceptor for leukocyte integrins: a novel pathway for inflammatory cell recruitment*, J. Exp. Med., 2003, 198, 1507– 1515.

Chen FE, Ghosh G, *Regulation of DNA binding by Rel/NF-KB transcription factors: structural views,* Oncogene, 1999, 18, 6845–6852.

Chong Curtis R, Xiaochun Chen, Lirong Shi, Jun O Liu1,David J Sullivan, Jr, *A clinical drug library screen identifies astemizole as an antimalarial agent*, Nature Chemical Biology, 2006, 2, 415-416.

Chong CR, Sullivan DJ Jr, *New uses for old drugs*, Nature, 2007, 448, 645-646.

Colombo G, Morra G, Meli M, Verkhivker G. *Understanding ligand-based modulation of the Hsp90 molecular chaperone dynamics at atomic resolution,* Proc.Natl.Acad.Sci.U.S.A., 2008, 105(23), 7976-81.

Coon JT, Ernst E, *Complementary and alternative therapies in the treatment of chronic hepatitis C: a systematic review*. J Hepatol., 2004, 40(3), 491-500.

Cordier F, Dingley AJ, Grzesiek S, *A doublet-separated sensitivity-enhanced HSQC for the determination of scalar and dipolar one-bond J-couplings*. J Biomol NMR., 1999, 13(2), 175-80.

Cornilescu G, Marquardt JL, Ottiger M, Bax A, *Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase,* J. Am. Chem. Soc, 1998, 120, 6836–6837.

CPMD V3.11, Copyright IBM Corporation 1990–2008; MPI für Festkörperforschung, Stuttgart, Germany, 1997–2001

Curioni A, Mordasini T, Andreoni W. *Enhancing the accuracy of virtual screening: molecular dynamics with quantum-refined force fields*, J Comput Aided Mol Des., 2004, 18(12), 773-84.

Dalvit C, Hadden DT, Sarver RW, Ho AM, Stockman BJ. *Multi-selective one dimensional proton NMR experiments for rapid screening and binding affinity measurements*. Comb. Chem. High Throughput Screen., 2003, 6(5), 445-53.

Daura X, Antes I, van Gunsteren WF, Thiel W, Mark AE, *The effect of motional averaging on the calculation of NMR-derived structural properties*, Proteins, 1999,

36(4), 542-55.

Dharmananda S, *REDUCING INFLAMMATION WITH DIET AND SUPPLEMENTS: The Story of Eicosanoid Inhibition*, http://www.itmonline.org/arts/lox.htm, Institute for Traditional Medicine, Portland, Oregon, 2003.

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*, J.Biomol.NMR., 1995, 6(3), 277-93.

DiMasi JA. *Rising research and development costs for new drugs in a cost containment environment.* Pharmacoeconomics, 1992, 1, Suppl. 1, 13-20

Di Stefano A, Caramori G, Oates T, Capelli A, Lusuardi M, Gnemmi I, Ioli F, Chung KF, Donner CF, Barnes PJ, Adcock IM, *Increased expression of nuclear factor-kappaB in bronchial biopsies from smokers and patients with COPD,* Eur.Respir.J., 2002, 20(3), 556-63.

Dominguez C, Boelens R, Bonvin AMJ, *HADDOCK: a protein-protein docking approach based on biochemical and/or biophysical information,* J.Am.Chem.Soc., 2003, 125, 1731-1737.

Dosset P, Marion D, Blackledge M, *Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data*, J.Biomol.NMR, 2000, 16, 23-28.

Dosset P, Hus JC, Marion D, Blackledge M, *A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings*, J Biomol NMR, 2001, 20(3), 223-31.

Eski M, Sahin I, Sengezer M, Serdar M, Ifran A. *Thalidomide decreases the plasma levels of IL-1 and TNF following burn injury: is it a new drug for modulation of systemic inflammatory response,* Burns, 2008, 34(1), 104-108.

Farrow NA, Muhandiram, R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE *Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation.* Biochemistry, 1993, 33, 5984–6003.

Ferrara P, Curioni A, Vangrevelinghe E, Meyer T, Mordasini T, Andreoni W, Acklin P, Jacoby E, *New scoring functions for virtual screening from molecular dynamics simulations with a quantum-refined force-field (QRFF-MD). Application to cyclin-dependent kinase 2,* J Chem Inf Model. 2006, 46(1), 254-63.

Fielding L, *NMR methods for the determination of protein-ligand dissociation*

constants, Curr Top Med Chem., 2003, 3(1), 39-53

Fujii T, Nakamura T, Iwamoto A, *Current concepts in SARS treatment*, J Infect Chemother., 2004, 10(1), 1-7.

Feng WY, *Metabolism of green tea catechins: an overview*, Curr Drug Metab, 2006, 7(7), 755-809

Funk CD, *Prostaglandins and leukotrienes: advances in eicosanoid biology*, Science, 2001 , 294(5548), 1871-5

Garrett DS, Powers R, Gronenborn AM, Clore GM, *A common sense approach to peak picking two-, three- and four-dimensional spectra using automatic computer analysis of contour diagrams*, J. Magn. Reson., 1991, 95, 214-220

Gayle MA, Slack JL, Bonnert TP, Renshaw BR, Sonoda G, Taguchi T, Testa JR, Dower SK, Sims JE, *Cloning of a putative ligand for the T1/ST2 receptor*, J.Biol.Chem., 1996, 271(10), 5784-9.

Germani A, Limana F, Capogrossi MC, *Pivotal advances: high-mobility group box 1 protein--a cytokine with a role in cardiac repair*, J Leukoc Biol., 2007, 1, 41-5.

Ghosh D, Erman M, Wawrzak Z, Duax WL, Pangborn W, *Mechanism of inhibition of 3 alpha, 20 beta-hydroxysteroid dehydrogenase by a licorice-derived steroidal inhibitor*, Structure. 1994, 2(10), 973-80.

Gijón MA, Leslie CC, *Regulation of arachidonic acid release and cytosolic phospholipase A2 activation*, J.Leukoc.Biol, 1999, 65(3), 330-6.

Gilmore TD, *Introduction to NF-kappaB: players, pathways, perspectives*, Oncogene, 2006, 25(51), 6680.

Gögelein H, Hüby A. *Interaction of saponin and digitonin with black lipid membranes and lipid monolayers*, Biochim.Biophys.Acta, 1984, 773(1), 32-38.

Goel A, Boland CR, Chauhan DP, *Specific inhibition of cyclooxygenase-2 (COX-2) expression by dietary curcumin in HT-29 human colon cancer cells.*, Cancer Lett., 2001, 172, 111–118.

Gouras GK, Beal MF, *Metal chelator decreases Alzheimer beta-amyloid plaques*, Neuron, 2001, 30(3), 641-2.

Grzesiek S, Bax A, *Improved 3D Triple-Resonance NMR Techniques Applied to a 31-kDa Protein,* Journal of Magnetic Resonance, 1992, 96(2), 432-440.

Grzesiek S, Bax A, *The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements*, J. Am. Chem. Soc., 1993, 115(26) , 12593-12594.

Guerois R, Nielsen JE, Serrano L, *Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations,* J.Mol.Biol, 2002, 320, 369-87.

Guvench O, MacKerell AD Jr, *Comparison of protein force fields for molecular dynamics simulations,* Methods Mol.Biol., 2008, 443, 63-88

Hardman CH, Broadhurst RW, Raine AR, Grasser KD, Thomas JO, Laue ED, *Structure of the A-domain of HMG1 and its interaction with DNA as studied by heteronuclear three- and four-dimensional NMR spectroscopy*, Biochemistry, 1995, 34(51),16596-607.

Hayden MS, Ghosh S, *Signaling to NF-kappaB*, Genes Dev. 18, 2195–2224 (2004)

Hayashi H, Huang P, Takada S, Obinata M, Inoue K, Shibuya M, Ebizuka Y, *Differential expression of three oxidosqualene cyclase mRNAs in Glycyrrhiza glabra,* Biol Pharm Bull, 27(7), 1086-92.

Hegmann T, *A tale of two coxibs,* JAAPA, 2005, 18, 14–17.

Hess B, Bekker H, Berendsen HJC, Fraaije J, *LINCS: A linear constraint solver for molecular simulations*, J.Comput.Chem., 1997, 18, 1463.

Huttunen HJ, Fages C, Kuja-Panula J, Ridley AJ, Rauvala H. *Receptor for advanced glycation end products-binding COOH-terminal motif of amphoterin inhibits invasive migration and metastasis,* Cancer Res., 2002, 62(16), 4805-11.

Hwang TL, Shaka AJ, *Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients excitation sculpting,* J.Magn.Res. Series A, 1995, 112(2), 275-279.

Jin SH, Kim TI, Han DS, Shin SK, Kim WH, *Thalidomide suppresses the interleukin 1beta-induced NFkappaB signaling pathway in colon cancer cells,* Ann.N.Y.Acad.Sci., 2002, 973, 414-8.

Kaitin KI, DiCerbo PA, Lasagna L, *The new drug approvals of 1987, 1988 and 1989: trends in drug development.* J Clin Pharmacol 1991; 31 (2): 116-22

Kaitin KI, Bryant NR, Lasagna L, *The role of the research based pharmaceutical*

industry in medical progress in the United States. J Clin Pharmacol 1993; 33 (5), 412-7

Khanapure SP, Garvey DS, Janero DR, Letts LG, *Eicosanoids in inflammation: biosynthesis, pharmacology, and therapeutic frontiers,* Curr.Top.Med.Chem., 2007, 7(3), 311-40.

Kim SW, Rhee HJ, Ko J, *Inhibition of cytosolic phospholipase A2 by annexin I: specific interaction model and mapping of the interaction site,* J.Biol.Chem., 2001, 276, 15712-1571.

Kokkola R, Li J, Sundberg E, Aveberger AC, Palmblad K, Yang H, Tracey KJ, Andersson U, Harris HE. *Successful treatment of collagen-induced arthritis in mice and rats by targeting extracellular high mobility group box chromosomal protein 1 activity,* Arthritis Rheum., 2003, 48(7), 2052-8.

Komai-Koma M, Xu D, Li Y, McKenzie AN, McInnes IB, Liew FY, *IL-33 is a chemoattractant for human Th2 cells,* Eur J Immunol, 2007, 37(10), 2779-86.

Kumar S, Minnich MD, Young PR, *ST2/T1 protein functionally binds to two secreted proteins from Balb/c 3T3 and human umbilical vein endothelial cells but does not bind interleukin* 1. J Biol Chem. 1995, 270(46), 27905-13.

Kundu JK, Surh YJ, *Breaking the relay in deregulated cellular signal transduction as a rationale for chemoprevention with anti-inflammatory phytochemicals,* Mutat.Res. 2005 , 591(1-2), 123-46.

Hayden MS, Ghosh S, *Signaling to NF-kappaB.* Genes Dev., 2004, 18(18), 2195-224.

Hayden MS, Ghosh S, *Shared principles in NF-kappaB signaling*, Cell, 2008, 132(3), 344-62.

Ichiyama T, Hasegawa S, Matsubara T, Hayashi T, Furukawa S, *Theophylline inhibits NF-kappa B activation and I kappa B alpha degradation in human pulmonary epithelial cells,* Naunyn Schmiedebergs Arch Pharmacol., 2001, 364(6), 558-61.

Ismaili N, Garabedian MJ, *Modulation of glucocorticoid receptor function via phosphorylation.* Ann N Y Acad Sci 2004;1024, 86-101.

Ito K, Caramori G, Lim S, Oates T, Chung KF, Barnes PJ, Adcock IM, *Expression and activity of histone deacetylases in human asthmatic airways*, Am.J.Respir.Crit.Care Med., 2002, 166(3), 392-6.

Ito K, Lim S, Caramori G, Cosio B, Chung KF, Adcock IM, Barnes PJ, *A molecular mechanism of action of theophylline: Induction of histone deacetylase activity to*

decrease inflammatory gene expression, Proc.Natl.Acad.Sci.U.S.A., 2002, 25, 99(13), 8921-6.

Johnson BA, Blevins RA, *NMRView: A computer program for the visualization and analysis of NMR data, Journal of Biomolecular NMR*, 1994, 4, 603-614.

Le Drean Y, Mincheneau N, Le Goff P, Michel D, *Potentiation of glucocorticoid receptor transcriptional activity by sumoylation*, Endocrinology, 2002, 143, 3482-3489.

Li J, Kokkola R, Tabibzadeh S, Yang R, Ochani M, Qiang X, Harris HE, Czura CJ, Wang H, Ulloa L, Wang H, Warren HS, Moldawer LL, Fink MP, Andersson U, Tracey KJ, Yang H, *Structural basis for the proinflammatory cytokine activity of High Mobility Group Box 1,* Mol.Med., 2003, 9(1-2), 37-45.

He Q, Liang CH, Lippard SJ, *Steroid hormones induce HMG1 overexpression and sensitize breast cancer cells to cisplatin and carboplatin,* Proc.Natl.Acad.Sci.U.S.A, 2000, 97, 5768–5772.

Holmstrom S, Van Antwerp ME, Iniguez-Lluhi JA, *Direct and distinguishable inhibitory roles for SUMO isoforms in the control of transcriptional synergy,* Proc.Natl.Acad.Sci.U.S.A., 2003, 100, 15758-15763.

Hu HY, Horton JK, Gryk MR, Prasad R, Naron JM, Sun DA, Hecht SM, Wilson SH, Mullen GP, *Identification of small molecule synthetic inhibitors of DNA polymerase beta by NMR chemical shift mapping.* J. Biol. Chem., 2004, 279, 39736–39744.

Humphrey W, Dalke A, Schulten K, *VMD - Visual Molecular Dynamics*, J.Molec.Graphics, 1996, 14, 33-38.

Kim YS, Ahn Y, Hong MH, Joo SY, Kim KH, Sohn IS, Park HW, Hong YJ, Kim JH, Kim W, Jeong MH, Cho JG, Park JC, Kang JC, *Curcumin attenuates inflammatory responses of TNF-alpha-stimulated human endothelial cells,* J.Cardiovasc.Pharmacol. 2007, 50(1), 41-9.

Knapp S, Müller S, Digilio G, Bonaldi T, Bianchi ME, Musco G, T*he long acidic tail of high mobility group box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes,* Biochemistry, 2004, 43(38), 11992-7.

Li G, Larson EB, Sonnen JA, Shofer JB, Petrie EC, Schantz A, Peskind ER, Raskind MA, Breitner JC, Montine TJ. *Statin therapy is associated with reduced neuropathologic changes of Alzheimer disease.* Neurology, 2007 ;69(9), 878-85

Limana F, Germani A, Zacheo A, Kajstura J, Di Carlo A, Borsellino G, Leoni O,

Palumbo R, Battistini L, Rastaldo R, Müller S, Pompilio G, Anversa P, Bianchi ME, Capogrossi MC, *Exogenous high-mobility group box 1 protein induces myocardial regeneration after infarction via enhanced cardiac C-kit+ cell proliferation and differentiation.* Circ Res. 2005, 97(8), 73-83.

Lykissas MG, Beris AE, *The role of erythropoietin as an inhibitor of tissue ischemia,* Int.J.Biol.Sci., 2008, 4(3), 161-8.

Lukacova V, Balaz S, *Multimode ligand binding in receptor site modeling: implementation in CoMFA.* J Chem Inf Comput Sci. 2003 -;43(6), 2093-105

Ludwiczek S, Theurl I, Muckenthaler MU, Jakab M, Mair SM, Theurl M, Kiss J, Paulmichl M, Hentze MW, Ritter M, Weiss G, *Ca2+ channel blockers reverse iron overload by a novel mechanism via divalent metal transporter-1,* Nature Medicine, 2007, 13(4), 448-54.

Maroon JC, Bost JW, Borden MK, Lorenz KM, Ross NA, *Natural antiinflammatory agents for pain relief in athletes,* Neurosurg Focus., 2006, 21(4), 11

Matsui S, Matsumoto H, Sonoda Y, Ando K, Aizu-Yokota E, Sato T, Kasahara T, *Glycyrrhizin and related compounds down-regulate production of inflammatory che, mokines IL-8 and eotaxin 1 in a human lung fibroblast cell line.* Int Immunopharmacol. 2004, 4(13)1633-44.

Mayer M, Meyer B, *Characterization of ligand binding by saturation transfer difference NMR spectroscopy,* Angew.Chem.Int.Ed.Engl., 1998, 38, 1784-1788.

McLachlan AD, *Rapid comparison of protein structures.* Acta Crystallogr. A, 1982, 38, 871–873.

Meierhofer C, Dunzendorfer S, Wiedermann CJ. *Theoretical basis for the activity of thalidomide,* BioDrugs, 2001, 15, 681-703.

Meiler J, Blomberg N, Nilges M, Griesinger C, *A new approach for applying residual dipolar couplings as restraints in structure elucidation,* J.Biomol.NMR, 2000, 17, 185.

Miller AM, Xu D, Asquith DL, Denby L, Li Y, Sattar N, Baker AH, McInnes IB, Liew FY. *IL-33 reduces the development of atherosclerosis,* J.Exp.Med., 2008, 205(2), 339-46.

Mitsutake A, Sugita Y, Okamoto Y. *Generalized-ensemble algorithms for molecular simulations of biopolymers,* Biopolymers, 2001, 60(2), 96-123.

Moncado S, Gryglewski R, Bunting S, Vane JR, *An enzyme isolated from arteries*

transforms prostaglandin endoperoxides to an unstable substance that inhibits platelet aggregation, Nature, 1976, 263, 663–665.

Moreira AL, Friedlander DR, Shif B, Kaplan G, Zagzag D. *Thalidomide and a thalidomide analogue inhibit endothelial cell proliferation in vitro,* J Neurooncol. 1999, 43(2), 109-14.

Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ, *Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function* J. Computational Chemistry, 1998, 19, 1639-1662.

Müller S, Ronfani L, Bianchi ME. *Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function,* J.Intern.Med., 2004, 255(3), 332-43.

Müller S, Scaffidi P, Degryse B, Bonaldi T, Ronfani L, Agresti A, Beltrame M, Bianchi ME, *New EMBO members' review: the double life of HMGB1 chromatin protein: architectural factor and extracellular signal,* EMBO J., 2001, 20(16), 4337-40.

Murakami A, Ohigashi H.Targeting NOX, *INOS and COX-2 in inflammatory cells: chemoprevention using food phytochemicals,* Int.J.Cancer, 2007, , 121(11), 2357-63.

D. Neuhaus, M.P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis,* John Wiley & Sons, New York, 2000.

Ogata H, Hibi T. *Cytokine and anti-cytokine therapies for inflammatory bowel disease,* Curr.Pharm. Des., 2003,9(14), 1107-13.

Ohndorf UM, Rould MA, He Q, Pabo CO, Lippard SJ, *Basis for recognition of cisplatin-modified DNA by highmobility- group proteins*, Nature, 1999, 399, 708–712.

J. Cavanagh, A. G. Palmer III, P. E. Wright, M. Rance, *Sensivity improvement in proton two-dimensional heteronuclear relay spectroscopy,* J. Magn. Reson. 1991, 91, 429

Palumbo R, Sampaolesi M, De Marchis F, Tonlorenzi R, Colombetti S, Mondino A, Cossu G, Bianchi ME, *Extracellular HMGB1, a signal of tissue damage, induces mesoangioblast migration and proliferation,* J.Cell.Biol., 2004,164(3), 441-9.

Perkins ND, *Post-translational modifications regulating the activity and function of the nuclear factor kappa B pathway,* Oncogene, 2006, 25(51), 6717-30.

Perola E, Walters WP, Charifson PS, *A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance,* Proteins, 2004, 56(2), 235-49

Piotto M, Saudek V, Sklenár V. *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions,* J.Biomol.NMR., 1992, 2(6), 661-5.

Ploeger, B, Mensinga, T, Sips, A, Seinen, W, Meulenbelt, J, DeJongh, J, *The pharmacokinetics of glycyrrhizic acid evaluated by physiologically based pharmacokinetic modeling.* Drug Metab. Rev. 33, 125–147.

Plummer SM, Holloway KA, Manson MM, Munks RJ, Kaptein A, Farrow S, Howells L, *Inhibition of cyclo-oxygenase 2 expression in colon cells by the chemopreventive agent curcumin involves inhibition of NF-kappaB activation via the NIK/IKK signalling complex.* Oncogene, 1999, 18, 6013–6020.

Pöhler JR, Lilley DM, *The interaction of HMG-Box proteins with the four-way DNA junction,* Biochem. Soc.Trans., 1997, 25(4), S647

Porto A, Palumbo R, Pieroni M, Aprigliano G, Chiesa R, Sanvito F, Maseri A, Bianchi ME, *Smooth muscle cells in human atherosclerotic plaques secrete and proliferate in response to high mobility group box 1 protein*, FASEB J., 2006, 20(14), 2565-6

Rauvala H, *Interview with Dr. Heikki Rauvala regarding pivotal advance: analysis of proinflammatory activity of highly purified eukaryotic recombinant HMGB1 (amphoterin). Interview by Marco E Bianchi*, J.Leukoc.Biol., 2007, 81(1), 46-8.

Rawlins MD, *Cutting the cost of drug development?*, Nat.Rev.Drug.Discov., 2004, 3(4), 360-4.

Read CM, Cary PD, Crane-Robinson C, Driscoll PC, Norman DG, *Solution structure of a DNA-binding domain from HMG1*, Nucleic Acids Res., 1993, 21(15), 3427-36.

Rhen T, Cidlowski J.A. *Antiinflammatory action of glucocorticoids--new mechanisms for old drugs,* 2005, N.Engl.J.Med., 353(16), 1711-23.

Richardson P, Hideshima T, Anderson K, *Thalidomide: emerging role in cancer medicine,* Annu.Rev. Med., 2002, 53, 629-57.

Rosen J, Miner JN, *The search for safer glucocorticoid receptor ligands*, Endocr.Rev., 2005, 26(3), 452-64.

Samuelsson G, *Drugs of natural origin – a textbook of pharmacognosy,* 4th revised edition, potekansocieteten, Swedish Pharmaceutical Press, Stockholm, 1999.

Sandeep TC, Yau JL, MacLullich AM, Noble J, Deary IJ, Walker BR, Seckl JR, *11Beta-hydroxysteroid dehydrogenase inhibition improves cognitive function in healthy elderly men and type 2 diabetics*, Proc. Natl.Acad.Sci.U.S.A.,2004, 101(17), 6734-9.

Saiah E, *The role of 11beta-hydroxysteroid dehydrogenase in metabolic disease and therapeutic potential of 11beta-hsd1 inhibitors,* Curr.Med.Chem., 2008, 15(7), 642-9.

Sakamoto R, Okano M, Takena H, Ohtsuki K. *Inhibitory effect of glycyrrhizin on the phosphorylation and DNA-binding abilities of high mobility group proteins 1 and 2 in vitro,* Biol.Pharm.Bull., 2001, 24(8), 906-11.

Sappington PL, Yang R, Yang H, Tracey KJ, Delude RL, Fink MP, *HMGB1 B box increases the permeability of Caco-2 enterocytic monolayers and impairs intestinal barrier function in mice,* Gastroenterology, 2002, 123(3), 790-802.

Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S. *An easy way to include weak alignment constraints into NMR structure calculations.* J Biomol NMR. 2001, 21(3), 275-80.

Sato E, Nelson DK, Koyama S, Hoyt JC, Robbins RA. *Inflammatory cytokines modulate eotaxin release by human lung fibroblast cell line,* Exp.Lung.Res., 2001, 27(2), 173-83.

Sattler M, Schleucher J, Griesinger C, *Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients*, Prog.NMR Spectrosc., 1999, 34, 93-158.

Scaffidi P, Misteli T, Bianchi ME, *Release of chromatin protein HMGB1 by necrotic cells triggers inflammation.* Nature, 2002, 418(6894), 191-5.

Schacke H, Docke WD, Asadullah K. *Mechanisms involved in the side effects of glucocorticoids.* Pharmacol Ther 2002;96, 23-43.

Schieborr U, Vogtherr V, Elshorst B, Betz M, Grimme S, Pescatore B, Langer T, Saxena K, Schwalbe H, *How Much NMR Data Is Required To Determine a Protein-Ligand Complex Structure?*, ChemBioChem, 2005, 6, 1-9.

Schmitz J, Owyang A, Oldham E, Song Y, Murphy E, McClanahan TK, Zurawski G, Moshrefi M, Qin J, Li X, Gorman DM, Bazan JF, Kastelein RA. *IL-33, an interleukin-1-like cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokines.* Immunity. 2005 ;23(5), 479-90.

Schuettenkolpf AW, van Aalten DMF, *PRODRG - a tool for high-throughput crystallography of protein-ligand complexes*, Acta Crystallographica, 2004, D60, 1355-1363.

Seibert K, Masferrer JL, *Role of inducible cyclooxygenase (COX-2) in inflammation,*

Receptor, 1994, 4, 17–23.

Shibata S, *A drug over the millennia: pharmacognosy, chemistry, and pharmacology of licorice,* Yakugaku Zasshi, 2000, 120(10), 849-62.

Sitia G, Iannacone M, Müller S, Bianchi ME, Guidotti LG, *Treatment with HMGB1 inhibitors diminishes CTL-induced liver disease in HBV transgenic mice,* J.Leukoc.Biol., 2007, 81(1), 100-7.

Strimpakos AS, Sharma RA. *Curcumin: preventive and therapeutic properties in laboratory studies and clinical trials,* Antioxid.Redox Signal., 2008, 10(3), 511-45.

Taniguchi N, Kawahara K, Yone K, Hashiguchi M, Goto M, *High mobility group box chromosomal protein 1 plays a role in the pathogenesis of rheumatoid arthritis as a novel cytokine,* Arthritis Rheum. 48, 971–981, 2003.

Tianwei T, Qing H, Qiang1 L, *Purification of glycyrrhizin from Glycyrrhiza uralensis Fisch with ethanol/phosphate aqueous two phase system* , Biotechnology Letters 2002, 24, 1417-1420.

Tjandra N, Wingfield P, Stahl S, Bax A, *Anisotropic rotational diffusion of perdeuterated HIV protease from 15N NMR relaxation measurements at two magnetic fields,* J.Biomol.NMR, 1996, 8(3), 273-84.

Tomita K, Chikumi H, Tokuyasu H, Yajima H, Hitsuda Y, Matsumoto Y, Sasaki T. *Functional assay of NF-kappaB translocation into nuclei by laser scanning cytometry: inhibitory effect by dexamethasone or theophylline,* Naunyn Schmiedebergs Arch Pharmacol. 1999,

Troullier N, Martins JL, *Efficient pseudopotentials for plane-wave calculations,* Phys.Rev.B Condens. Matter, 1991, 43(3), 1993-2006.

Ueda T, Shirakawa H, Yoshida M, *Involvement of HMGB1 and HMGB2 proteins in exogenous DNA integration reaction into the genome of HeLa S3 cells.* Biochim. Biophys. Acta, 2002, 1593, 77–84.

Ullian ME. *The role of corticosteroids in the regulation of vascular tone,* Cardiovasc.Res. ,1999, 41, 55-64.

Ulloa L, Ochani M, Yang H, Tanovic M, Halperin D, Yang R, Czura CJ, Fink MP, Tracey KJ. *Ethyl pyruvate prevents lethality in mice with established lethal sepsis and systemic inflammation,* Proc.Natl.Acad.Sci.U.S.A., 2002, 99(19), 12351-6.

Urnov FD, Wolffe AP, *Chromatin remodeling and transcriptional activation: the cast*

(in order of appearance), Oncogene, 2001, 20(24), 2991-3006.

van der Eerden BC, Karperien M, Wit JM. *Systemic and local regulation of the growth plate.* Endocr Rev 2003;24, 782-801.

van Dijk ADJ, Fushman D, Bonvin AMJJ, *Various strategies of using residual dipolar couplings in NMR-driven protein docking: Application to Lys48-linked di-ubiquitin and validation against 15N-relaxation data*, Proteins: Struc. Funct. & Bioinformatics, 2005, 60, 367-381.

van Dijk ADJ, Kaptein R, Boelens R, Bonvin AMJJ, *Combining NMR relaxation with chemical shift perturbation data to drive protein-protein docking,* J. Biomol. NMR, 2006, 34, 237-244.

Vane JR, Botting RM. *A better understanding of anti-inflammatory drugs based on isoforms of cyclooxygenase (COX-1 and COX-2),* Adv.Prostaglandin Thromboxane Leukot.Res., 1995, 23, 41-8.

Vane J. *The mechanism of action of anti-inflammatory drugs.* Int J Clin Pract Suppl. 2003, (135), 2

van Gunsteren WF, Billeter SR, Eising AA, Hunenberger PH, Kruger P, Mark AE, Scott WRP, Tironi IG, *Biomolecular Simulation: The GROMOS96 manual and user guide.* Zurich, Switzerland: Hochschulverlag AG an der ETH Zurich. 1996.

van Rossum TG, Vulto AG, Hop WC, Schalm SW, *Glycyrrhizin-induced reduction of ALT in European patients with chronic hepatitis* C, Am.J.Gastroenterol., 2001, 96(8), 2432-7.

Wang H, Bloom O, Zhang M, Vishnubhakat JM, Ombrellino M, Che J, Frazier A, Yang H, Ivanova S, Borovikova L, Manogue KR, Faist E, Abraham E, Andersson J, Andersson U, Molina PE, Abumrad NN, Sama A, Tracey KJ. *HMG-1 as a late mediator of endotoxin lethality in mice.* Science, 1999, 85(5425), 248-51.

Weir HM, Kraulis PJ, Hill CS, Raine AR, Laue ED, Thomas JO, *Structure of the HMG box motif in the B-domain of HMG1,* EMBO J.,1993, 12(4), 1311-9.

Wheeler DS, Catravas JD, Odoms K, Denenberg A, Malhotra V, Wong HR, *Epigallocatechin-3-gallate, a green tea-derived polyphenol, inhibits IL-1 beta-dependent proinflammatory signal transduction in cultured respiratory epithelial cells,* J.Nutr., 2004, 134(5), 1039-44.

Willker W, Stelten J, Leibfritz D, *A highly Selective Proton-detected 1D H,X Correlation,* J.Magn.Reson.A, 1994, 107, 94-98.

Wong CK, Lam CW, Wu AK, Ip WK, Lee NL, Chan IH, Lit LC, Hui DS, Chan MH, Chung SS, Sung JJ. *Plasma inflammatory cytokines and chemokines in severe acute respiratory syndrome,* Clin.Exp.Immunol., 2004, 136(1), 95-103.

R.J. Woods, R.A. Dwek, C.J. Edge, B. Fraser-Reid. 1995. *Molecular mechanical and molecular dynamical simulations of glycoproteins and oligosaccharides. 1. GLYCAM_93 parameter development.* J. Phys. Chem. 99, 3832-3846.

Wittekind M, Mueller L, *HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins,* Magn.Reson.B, 1993, 101, 201-205.

Yang H, Ochani M, Li J, Qiang X, Tanovic M, Harris HE, Susarla SM, Ulloa L, Wang H, DiRaimo R, Czura CJ, Wang H, Roth J, Warren HS, Fink MP, Fenton MJ, Andersson U, Tracey KJ. *Reversing established sepsis with antagonists of endogenous high-mobility group box 1.* Proc Natl Acad Sci U S A., 2004, 101(1), 296-301.

Yang H, Wang H, Czura CJ, Tracey KJ, *The cytokine activity of HMGB1,* J Leukoc Biol., 2005, 78(1), 1-8.

Yu M, Wang H, Ding A, Golenbock DT, Latz E, Czura CJ, Fenton MJ, Tracey KJ, Yang H. HMGB1 signals through toll-like receptor (TLR) 4 and TLR2. Shock. 2006 Aug;26(2), 174-9

Zagrovic B, van Gunsteren WF, *Comparing atomistic simulation data with the NMR experiment: how much can NOEs actually tell us?,* Proteins: Struct. Funct. Bioinform., 2006, 63, 210.

Zhong H, May MJ, Jimi E, Ghosh S, *The phosphorylation status of nuclear NF-kappa B determines its association with CBP/p300 or HDAC-1*, Mol Cell., 2002, 9(3), 625-36.